

Russell, Daniel John (2013) Lycopene carotenogenesis and function in the haloarchaeon *Haloferax volcanii*. PhD thesis, University of Nottingham.

**Access from the University of Nottingham repository:**  
<http://eprints.nottingham.ac.uk/27907/1/602481.pdf>

**Copyright and reuse:**

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

- Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners.
- To the extent reasonable and practicable the material made available in Nottingham ePrints has been checked for eligibility before being made available.
- Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
- Quotations or similar reproductions must be sufficiently acknowledged.

Please see our full end user licence at:  
[http://eprints.nottingham.ac.uk/end\\_user\\_agreement.pdf](http://eprints.nottingham.ac.uk/end_user_agreement.pdf)

**A note on versions:**

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact [eprints@nottingham.ac.uk](mailto:eprints@nottingham.ac.uk)

**Lycopene Carotenogenesis and Function in the  
Haloarchaeon *Haloferax volcanii***

**Daniel Russell (Bsc.)**

**Thesis submitted to the University of Nottingham for  
the degree of Doctor of Philosophy, September 2010**

**Supervisor**

**David J Scott**

**School of Biosciences**

**National Centre for Macromolecular Hydrodynamics**

**Sutton Bonington**

**University of Nottingham**

**UNITED KINGDOM**

**Contents**

Contents .....I

    List of Figures .....iv

    List of Tables.....viii

    Abbreviations .....ix

    Abstract ..... x

    Acknowledgements.....xii

Chapter 1 Introduction ..... 1

    1.1 Archaea, The Domain ..... 1

    1.2 Differences Between Archaea and the Other Domains..... 5

    1.3 Division of the Archaeal Domain ..... 10

    1.4 Crenarchaeota ..... 13

    1.5 Nanoarchaeota..... 14

    1.6 Korarchaeota ..... 15

    1.7 Euryarchaeota ..... 16

    1.8 Archaeal Physiology ..... 17

    1.9 Archaeal Habitats ..... 17

    1.10 Halophilic archaea..... 19

    1.11 *Haloferax volcanii*..... 20

    1.12 Carotenoids ..... 23

    1.13 Carotenoids in animals..... 23

    1.14 Carotenoids in plants..... 25

1.15 Carotenoids in bacteria and protists .....	25
1.16 Carotenoids in archaea .....	26
1.17 Carotenoids and Reactive Oxygen Species.....	27
1.19 Archaeal Carotenogenesis .....	30
1.20 Potential Applications of This Research .....	31
Chapter 2 Materials and Methods .....	33
2.1 Archaeal Strains .....	33
2.2 Bacterial Strains .....	33
2.3 Plasmids and Vectors .....	35
2.4 Luria Bertani Media .....	36
2.5 Antibiotics .....	37
2.6 30% Salt Water.....	37
2.7 10x YPC Stock.....	38
2.8 HvYPC Media.....	38
2.9 Hv Na-Lac Media .....	39
2.10 Growth of <i>H. volcanii</i> Strains .....	40
2.11 Harvesting of <i>H. volcanii</i> Cultures .....	40
2.12 Extraction of Lycopene .....	40
2.13 HPLC Separation of <i>H. volcanii</i> Extracts .....	41
2.14 Polymerase Chain Reactions (PCRs).....	42
2.15 <i>H. volcanii</i> Genomic DNA Extraction .....	43
2.16 Digestion of PCR Products and Plasmid Vectors .....	43



2.17 Ligation of Digested PCR Products and Plasmid Vectors for Cloning	44
2.18 Electroporation of electrocompetent <i>E. coli</i> strains .....	44
2.19 DNA Agarose Gel Electrophoresis.....	47
2.20 Recombinant Protein Expression.....	47
2.21 Poly Arcylamide Gel Electrophoresis.....	48
Chapter 3 Bioinformatic Analysis.....	49
3.1 Introduction.....	49
3.2 Carotenoid Biosynthesis Pathway Elucidation .....	50
3.3 Protein Models .....	55
3.4 Discussion .....	62
Chapter 4 Lycopene Extraction and Quantification.....	67
4.1 Introduction.....	67
4.2 Lycopene Extraction .....	68
4.3 Lycopene Separation and Quantification .....	74
4.4 Discussion .....	85
Chapter 5: Changes in the Carotenoid Metabolite Profile of <i>Haloferax</i> <i>volcanii</i> Throughout Growth and Under Stress Conditions.....	88
5.1 Introduction.....	88
5.2 <i>Haloferax</i> Growth Under Non-Stress Conditions .....	89
5.3 Lycopene Production During Growth.....	95
5.4 Lycopene Production Under Stress Conditions .....	102
5.5 Chemically Defined Media Development .....	105
5.6 Carotenoid Production Under Oxidative Stress .....	108

5.7 Post Lycopene Carotenoid Products of <i>H. volcanii</i> .....	117
5.8 Discussion .....	125
Chapter 6 Construction of Gene Knockouts and Cloning for.....	130
Heterologous Protein Expression.....	130
6.1 Introduction.....	130
6.2 UV-Induced Gene Knockouts .....	134
6.3 Targeted Gene Knockouts .....	137
6.4 Cloning of crtI for Heterologous Protein Expression .....	148
6.5 Discussion .....	153
Chapter 7 The Carotenoid Metabolite Profile of <i>H. volcanii</i> Mutants and Their Response to Stress Conditions .....	158
7.1 Introduction .....	158
7.2 Lycopene Deficient Mutant <i>H. volcanii</i> Growth.....	159
7.3 $\Delta crtI$ Carotenoid Metabolite Profile .....	165
7.4 $\Delta crtI$ Survival Under Stress Conditions.....	166
7.5 Discussion .....	168
Chapter 8 Conclusions and Discussion.....	173
8.1 Major findings of this study.....	173
8.2 Carotenoid biosynthesis in <i>H. volcanii</i> .....	173
8.3 Model Proteins .....	176
8.4 Deletion mutant.....	177
8.5 Response to oxidative stress .....	180
8.6 Further study .....	181

8.7 Concluding Statement ..... 183

References..... 184

## List of Figures

Fig 1.1 Phylogenetic tree of life (Barton et al. 2007) .....	7
Fig 1.2 A comparison of Archaeal lipid membrane and non-Archaeal lipid membrane. ....	7
Fig 1.3 A Phylogenetic tree of domain Archaea(Gibraldo & Brochier 2009) .	13
Fig 1.4 Phylogenetic tree of euryarchaeota (Allers & Mevarech 2005) .....	17
Fig 1.5 <i>Haloferax volcanii</i> (Allers 2004).....	21
Fig 1.7 Spectrum of lycopene, with maxima at 447.6, 466.7 and 498.1 nm ...	23
Fig 3.1, The Presumptive Lycopene Biosynthesis Pathway in <i>H. volcanii</i> . ...	56
Fig 3.2 I-Tasser created model of CrtI, .....	58
Fig 3.3 I-Tasser created model of CrtB2, .....	59
Fig 3.4 I-Tasser created model of CrtI, .....	60
Fig 4.1, Variance in extraction yield with altered lysate:propan-2-ol volume.	75
Fig 4.2, The chemical structure of lycopene and $\beta$ -carotene .....	78
Fig 4.3, Total lycopene mass in the sample compared to lycopene absorbance (450nm) peak height/ $\beta$ -carotene absorbance (450nm) peak height with 0.1 mg/ml $\beta$ -carotene. ....	80
Fig 4.4 Extraction and Quantification of Lycopene from <i>H. volcanii</i> .....	82
Fig 4.5, A chromatographic trace of a <i>H. volcanii</i> extract taken from <i>H. volcanii</i> H26 .....	83
Fig 5.1a, <i>H. volcanii</i> growth in HvYPC. ....	91
Fig 5.1b, <i>H. volcanii</i> growth in HvYPC. ....	92

Fig 5.2a Optical Density in Au plotted against viable count in cfu/ml.....	95
Fig 5.2b, Fluorescence images taken using Bac-Light .....	95
Fig 5.3a, Lycopene levels ( $\mu\text{g}$ of lycopene present in 200 ml culture) against time since start of exponential phase. ....	98
Fig 5.3b Lycopene per viable cell ( $\mu\text{g}/\text{cfu}$ ) against time since start of exponential phase across exponential growth. Error bars show standard error from 4 replicates. ....	99
Fig 5.4 Spectrum of lycopene in 1:19 THF:MeOH showing absorbance against wavelength. This was obtained using the method for isocratic HPLC and a photodiode array detector as detailed in Chapter 2. ....	100
Fig 5.5, Average lycopene levels in mid exponential and stationary growth of <i>H. volcanii</i> shown as lycopene per 200 ml culture ( $\mu\text{g}$ ). Error bars shown are standard error from 10 replicates. ....	103
.....	109
Fig 5.6a, Growth curves post lag phase of <i>H. volcanii</i> DS70 in both Hv YPC and Hv Na-Lac .....	109
.....	110
Fig 5.6b, Growth curves post lag phase of all strains in Hv Na-Lac. Optical density is plotted against time after the start of exponential phase. Errors shown are standard error from four replicates. ....	110
.....	113
Fig 5.7, Change in lycopene levels after 100 $\mu\text{M}$ hydrogen peroxide treatment, showing relative lycopene against time since treatment. Error shown is	

standard error from 4 replicates. ....	113
Fig 5.8, Spectrum of phytoene in 1:19 THF:MeOH showing absorbance (AU) against wavelength (nm). ....	115
Fig 5.9a, Phytoene change relative to control against time after shock. Errors shown are standard error from 4 replicates. ....	117
Fig 5.10, Lycopene changes relative to untreated control against time after treatment. Errors shown are standard error from 4 replicates.....	119
Fig 5.11a, from unknown peak 1, elution time 249s showing absorbance (AU) against wavelength (nm) .....	121
Fig 5.11b, Spectrum from unknown peak 2, elution time 180s showing absorbance (AU) against wavelength (nm).....	122
Fig 5.11c, Spectrum of Bacterioruberin showing absorbance (AU) against wavelength (nm) .....	123
Fig 6.1. Chromatograms showing lycopene levels in wildtype DS70 (a) and white strains (b,c,d).....	131
Fig 6.2.Digestion of pTA131 with BamHI, KpnI and SpeI .....	140
Fig 6.3 A schematic representation of the production of a knockout vector for <i>crtI</i> .....	146
Fig 6.4 Knockout Procedure. ....	148
Fig 6.5. Confirmation of <i>crtI</i> knockout .....	149
Fig 6.6. Sequence alignment of predicted $\Delta$ <i>crtI</i> sequence containing <i>trpA</i> (highlighted) with sequencing data.....	150
Fig 6.7 Expression of <i>crtI</i> in B121 Star Rossettas. ....	153

Fig 7.1. Chromatogram of extract from <i>H. volcanii</i> $\Delta$ crtI culture .....	166
Fig 7.2a, Survival rates represented as a percentage surviving, 100% being taken as the survival in unstressed cells counted in the same way. Survival rate is plotted against salinity.....	168
Fig 7.2b, Survival rates represented as a percentage surviving, 100% being taken as the survival in unstressed cells counted in the same way. Survival rate is plotted against H <sub>2</sub> O <sub>2</sub> concentration.. ..	169

# List of Tables

Table 1. <i>E. coli</i> genotypes used in this study .....	33
Table 2.2 Primers, DMSO concentrations, annealing temperatures and extension times for PCR. ....	44
Table 2.3 PCR Cycling Parameters.....	45
Table 3.1, Sequences selected for BLAST interrogation. ....	51
Table 3.2a, Transmembrane helices predicted by TMPred for CrtI. ....	64
Table 3.2b, Transmembrane helices predicted by TMPred for CrtB2. ....	64
Table 4.1, Extraction Efficiencies by Method. ....	72
Table 4.2, Chemical properties of lycopene and $\beta$ -carotene.....	81
Table 5.1 Length of Lag Phase by Strain.....	93
Table 6.1, Extraction Yields of UV Induced <i>H. volcanii</i> Mutant Strains Compared to Wildtype. ....	134
Table 6.2 Optimal DMSO Concentrations. ....	139
Table 6.3. Transformation efficiency of ligated plasmids. ....	141



## Abbreviations

5-FOA	5-Fluoroorotic acid
BLAST	Basic local alignment search tool
bp	Base pairs
DMSO	Dimethylsulfoxide
DNA	Deoxyribosnucleic acid
dNTP	Deoxynucleotide triphosphate
EDTA	Ethylenediamine tetraacetic acid
EtOH	Ethanol
HPLC	High performance liquid chromatography
Hv Na-Lac	<i>Haloferax volcanii</i> sodium lactate chemically defined medium
HvYPC	<i>Haloferax volcanii</i> yeast peptone casamino acid medium
IPTG	Isopropyl- $\beta$ -D thiogalactopyranoside
kb	Kilobase pairs
LB	Luria-Bertani
M	Molar
MeOH	Methanol
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
Rcf	Relative centrifugal force
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
THF	Tetrahydrofuran

## Abstract

The halophilic archaeon *Haloferax volcanii* is a carotenogenic organism, native to the dead sea. Its carotenogenic pathway was characterised by BLAST searching techniques. The pathway shows that isoprenoid production in *H. volcanii* utilises the mevalonate pathway. The first two carotenogenic enzymes, phytoene synthase (CrtB2) and phytoene desaturase (CrtI) were modelled, and investigated, results indicating a possibility that they are membrane bound. A method for extraction, separation and quantification of the carotenoids of *H. volcanii* was developed. This approach used phase extraction of an aqueous and an organic layer, followed by HPLC separation of the extractant components and analysis by comparison of peak absorption at 450nm of the carotenoid being studied to a  $\beta$ -carotene internal standard. This method was used to assess lycopene and phytoene levels throughout growth of the organism and under stress conditions. It was found that lycopene and phytoene are both synthesised at elevated levels in stationary phase growth. Additionally lycopene and phytoene levels show a response to oxidative stress, though not to other stressors, being degraded then accumulating to levels higher than found under none stress conditions. In order to gain insight into the function of lycopene in the organism the gene *crtI* coding for phytoene desaturase was deleted. This mutant was found to have all carotenoid production eliminated from the cell. This suggests that the carotenogenic enzymes may form a complex associated with the membrane. The mutant was

subjected to stress conditions, the results showing it was highly susceptible to oxidative attack. This study suggests that the carotenoids synthesised by *H. volcanii* are vital to its defence against reactive oxygen species (ROS).

## Acknowledgements

I would like to thank my supervisors Dr. David Scott, Prof Charles Hodgeman and Dr. Robert Linforth for their invaluable assistance in the planning and execution of this project. Thankyou to all of the academic and research staff at in the Food Sciences department at Sutton Bonington.

Thank you to Dr. Thorsten Allers for supplying several strains of *Haloferax volcanii* as well as plasmids required for the knock-out procedure. I would also like to thank Prof Ian connerton for his assistance at several key junctures in this project.

This project could not have been completed without the generous funding and resources provided by my sponsors the BBSRC and the University of Nottingham. The trust implicit in such funding is greatly appreciated.

I would like to thank the many people at Sutton Bonington who have helped me both academically and personally throughout my PhD. Tom Batstone, Dr. James Button, Emily Boothroyd, Selina Clayton, Phil Glover, Dr. Karin Heurlier, Dr Jon Hobman, Dr. Prudence Mutowo, Kate Portman, Gareth Williams and Khairi Zainol, thank you all.

Finally I would like to thank my family. My mum, Karen, for regular phone calls and support, my dad, John, for always being willing to listen to me talk about science, Ruth for her belief in me, regardless of how frustrating I became and my gran, Audrey, for being my gran.

**This thesis is dedicated to my nephew Mehedi Sarker, whose existence has made me realise that I'm a grown up. Research is built upon the past and is a gift to the future, and for me there is no better representation of the future than you, Mehedi.**

# **Chapter 1**

## **Introduction**

## **Chapter 1 Introduction**

### **1.1 Archaea, The Domain**

The classification system used for all organisms has changed several times over the past three centuries (Wolf *et al.* 2002). As scientific techniques for identifying organisms have improved the characteristics by which they are classified have changed. Linnaeus first sort to classify all organisms in the 18<sup>th</sup> century by recourse to study of their physical structure (Linnaeus 1758). The Linnaean classification system separates organisms into two distinct groups, animals and plants.

The differences between these two kingdoms were determined at the anatomical level, therefore organisms which did not anatomically resemble either plants or animals could not be categorised by this method. In 1866 Haeckel, extending from work done by Leeuwenhoek (Leeuwenhoek 1675 & 1683), therefore proposed a restructuring of classification into three groups, animals, plants and protists (Baptiste *et al.* 2004). This third group, protists, consisted of microscopic single celled organisms, an anatomy that defied classification as either plant or animal.

This classification system remained in place for nearly a century until in 1956 Copeland recognised, based on morphological and cytological data (Copeland, 1956) that the microscopic organisms should be further divided, into protists and monera. The kingdom monera consisted of those organisms lacking a nucleus, all of which are microscopic.

Less than 15 years later in 1969 the classification system was to altered once more with the recognition of the great differences between plants and fungi (Whittaker 1969). This structure, based on biochemical and microscopic evidence rather than anatomical differences appeared robust, however its pre-eminence was quickly challenged as molecular biological evidence revealed distinct differences between monera, distinct (Woese 1977 & 1990). This evidence hinted at differences between two groups of monera as large as the difference between that kingdom and any of the others.

These organisms, discovered in 1977 by Dr Carl Woese at the University of Illinois and dubbed archaea were first distinguished from bacteria by ribosomal ribonucleic acid (rRNA) sequence comparison (Woese & Fox 1977). This comparison showed that there were similar differences between archaeal rRNA and that of bacteria as between archaea or bacteria and eukaryotes. The differences between the various kingdoms of eukaryotes,



fungi, plants, animals and protists, was not nearly so large (Olsen *et al.* 1986; Woese *et al.* 1990).

The use of rRNA, particularly 16S or 18S subunit rRNA as used in the Woese and Olsen studies is advantageous as these RNAs exhibit very high sequence conservation, over large expanses of sequence, across all three domains. This allows for their use as indicators of evolutionary distance to be very accurate (Woese 1990). These differences encouraged the classification of life into three domains, eukarya, bacteria and archaea, with distinct genetic differences between the three. The more similar kingdoms of the eukarya then became a secondary classification step, rather than being evident at the broadest classification level (Fig 1.1).

In recent years further challenge has been made to this system, with the suggestion that taxonomy should delineate groups into monophyletic units rearranges the eukaryotes, whilst leaving archaea and bacteria untouched (Simpson & Roger 2004). Whilst the initial concept of this study was largely accepted (Harper *et al.* 2005) questions have since been raised as to whether the model proposed does indeed split organisms into monophyletic groups, whilst it is largely agreed that the current model is incorrect there is disagreement over what should replace it (Doolittle 2007 & 2010)

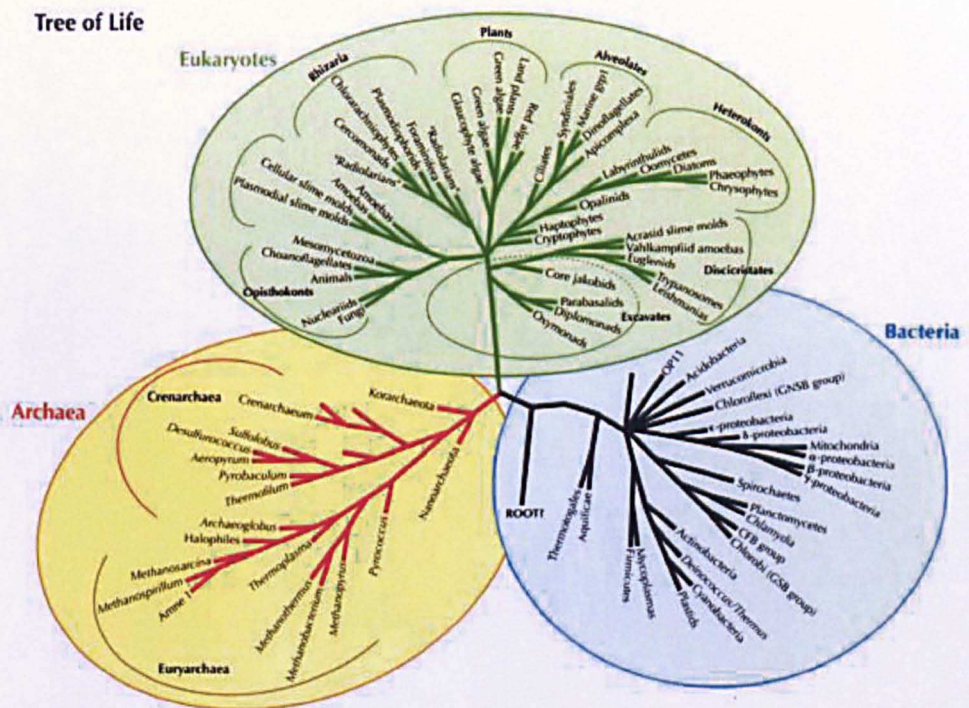


Fig 1.1 Phylogenetic tree of life (Barton *et al.* 2007)

## **1.2 Differences Between Archaea and the Other Domains**

Archaeal species exhibit characteristics at the molecular, biochemical and cellular levels that resemble both bacteria and eukaryotes, as well as having a host of features entirely distinct from either of the other domains (Keeling *et al.* 1994). The most striking differences that present themselves immediately on a careful analysis of archaea is their cell wall morphology and lipid composition, which is strikingly different to the other two domains (Siefert and Fox 1998).

At the cellular and sub-cellular level archaea initially resemble bacteria, as they, like bacteria, have circular chromosomes and contain plasmids (Bendich & Drlica 2000; Bernander 2000). Additionally there is no evidence of splicing, common among eukaryotes, in archaeal genes and their genes are sometimes organised into cistrons. Additionally they lack a nucleus or any true organelles, although unlike bacteria they possess some intracellular membrane structures, with organelle-like properties (Blaut *et al.* 1992). The lack of organelles leads to both bacteria and archaea having a lower internal complexity than eukaryotic cells (Sapp 2005).

Despite these initial similarities to bacteria, archaea show great similarity to

eukarya in their genetic reproduction machinery. RNA synthesis in archaea, for example makes use of TATA box binding proteins (TBPs) and an RNA polymerase (RNAP) similar to that of the eukaryotic RNAP II (Bell 2001; Cramer 2004)). Additionally the process of archaeal translation as a whole resembles that of eukarya (Woese 1998). The polyadenylation of RNAs commonly found in both eukarya and bacteria, however, does not occur in some archaea, such as *H. volcanii* (Slomovic *et al.* 2008). A further similarity to eukarya is the deoxyribonucleic acid (DNA) replication machinery, with initiation of DNA replication and replication itself being performed by homologs of the eukaryotic apparatus, including the DNA polymerase (Cann & Ishano 1999). Despite the similarities to eukaryotes in nucleic acid replication, including chromosomal replication, archaea show a bacteria like cell division process.

Archaeal features unique to themselves include, as noted earlier, the cell membrane. The lipids, which form the archaeal membrane, are typically formed of isoprenoids ether linked to glycerol-1-phosphate (Forterre *et al.* 2002), in contrast bacterial and eukaryotic lipids are formed from fatty acids ester linked to glycerol-3-phosphate (Fig 1.2). The cell wall, similarly is unlike bacteria or eukarya, whose cell walls are typically contain carbohydrate. Rather than being constructed of peptidoglycan, cellulose or chitin, archaeal cell walls are typically formed of the cells surface layer proteins (S-layer) with

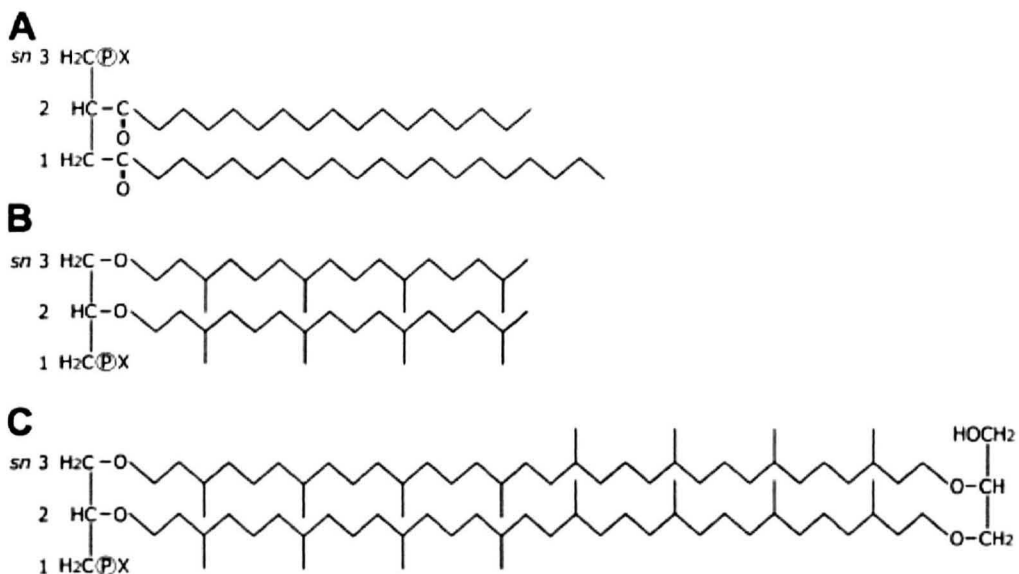


Fig 1.2 A comparison of Archaeal lipid membrane and non-Archaeal lipid membrane.

**A** shows a Bacterial or Eukaryal membrane, consisting of a linear fatty acid ester linked to a glycerol moiety. **B** shows a common Archaeal lipid, archaeol, consisting of isoprene chains with branching methyl groups ether linked to a glycerol moiety. **C** shows a tetraether lipid structure, unique to Archaea, in which the molecule crosses the entire membrane, with hydrophilic groups at both ends, such structures allow the formation of lipid monolayers. (Matsumi *et al.* 2010)

no additional carbohydrate structure (Keeling *et al.* 1994).

This unique blending of features has some notable effects, the most relevant of which is archaeal resistance to many antibiotics, due to the fact that most antibacterials target unique bacterial structures, such as peptidoglycan synthesis enzymes, peptidoglycan itself or the bacterial RNAP, none of which are present in archaea (Reeve 1999). Similarly any biocides used against eukaryotic organisms are typically targeted to specialist structures, to reduce their effects on unintended targets, thus many fungicides and herbicides target cell wall production, or other unique features, whilst pesticides may require an alkaline environment to become toxic. As archaea resemble eukarya and bacteria equally (Allers & Mevarech 2005) such targeted approaches are often ineffective against them, with the most general current biocides, such as nucleic acid chain terminators, being among the few effective choices.

Along with NACTs antibiotics which target the ribosome, particularly the large ribosomal subunit, such as chloramphenicol and linezolid are functional against the Archaea (Wilson 2011). There is also evidence that aminocoumarins and fluoroquinolones that target the DNA gyrase may be effective against archaea (Eckburg *et al.* 2003). These antibiotics have only been shown to be effective *in vivo*, and may be affecting bacterial populations that the archaea rely on rather than the archaea directly (Eckburg

*et al.* 2003). Further work performed on the methanogen *Methanobrevibacter smithii* showed that this organism was resistant to all antibiotics, except those which targeted both bacterial and eukaryotic populations as well as the archaeon (Dridi *et al.* 2011). This organism was resistant to ampicillin, streptomycin, gentamicin, rifampicin, ofloxacin, tetracycline and amphotericin B (MIC >100mg/ml). Additionally it was resistant to vancomycin (MIC>50mg/ml). Similar antibiotic profiles were shown for *Methanosphaera stadtmanae*, *Methanobrevibacter oralis* and *Methanomassiliicoccus luminyensis*. These strains were susceptible to bacitracin, which interferes with the dephosphorylation of isoprenyl pyrophosphate, the precursor to the isoprenoid pathway, nitroimidazoles, which affect nucleic acid synthesis, particularly under anaerobic conditions and squalamine (Dridi *et al.* 2011), whose precise mode of action is yet to be elucidated and may differ between organisms (Moore *et al.* 1993)

### **1.3 Division of the Archaeal Domain**

Beyond the division from bacteria archaea have since been subdivided into several groups, each of which has distinct differences from the others. These groups, or phylum, are the crenarchaeota, euryarchaeota, korarchaeota (Fox *et al.* 1977) and nanoarchaeota (Huber *et al.* 2003) (Fig 1.3).

The distinctions between the crenarchaeota, euryarchaeota and korarchaeota were based on 16S rRNA comparison in 1977 (Fox *et al.* 1977). The fourth phylum, nanoarchaeota was recognised as a separate group in 2003 upon the discovery of symbiotic archaea (Huber *et al.* 2003). Recently there has been a fifth archaeal group proposed, thaumarchaeota, comprising mesophilic crenarchaeota (Brochier-Armanet *et al.* 2008). Currently as more members of each archaeal group are discovered it could induce further changes in the classification of these organisms (Brochier-Armanet *et al.* 2008).

Currently the distinction between archaeal groups lies largely along lines defined by their metabolic biochemistry linked to nutrition (Falb *et al.* 2008). Whilst these divisions are also reinforced by molecular data, it may be that further investigation leads to a rearrangement of the archaeal phyla.

Such investigation includes the role that horizontal gene transfer plays in the definition of a species. There are currently a number of views on the relevance of this phenomenon to phylogeny. As most microbes, including Archaea, are capable of transferring genes from one species to another it has been suggested that the concept of microbial species, or indeed a tree of life is not relevant to them at all (Baptiste & Boucher 2008). It is suggested that a web or bush of life might be more appropriate (Doolittle & Baptiste 2007). This concept,



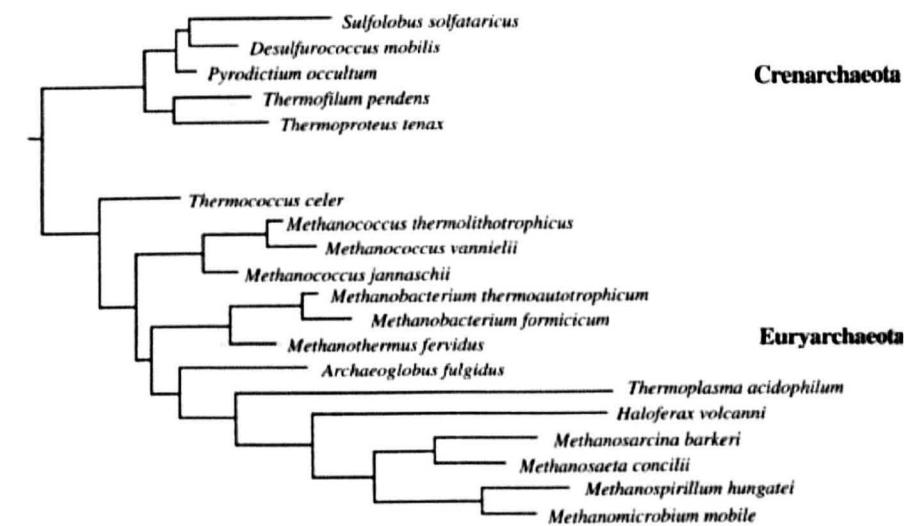
however looks only at the similarity of an organisms genome and, sometimes, plasmids and fails to address the very real descent by division, with daughter cells produced from a dividing ancestor. From this point of view a tree is still an accurate representation and easier to integrate into a global schema including Eukaryotes (Gribaldo & Brochier 2009).

A third view stands that the use of webs, or bushes, to represent the homologous gene transfer and the use of these to inform common descent, as acquired genes are transferred horizontally subsequently, can be used to produce accurate trees that take account of horizontal gene transfer as well as gene duplication, deletion and mutation (Gibraldo & Brochier 2009). This concept works well in conjunction with construction of a phylogenetic tree using a core conserved subset of genes, many of which are involved in nucleic acid processing, synthesis and degradation (Brochier *et al.* 2005).

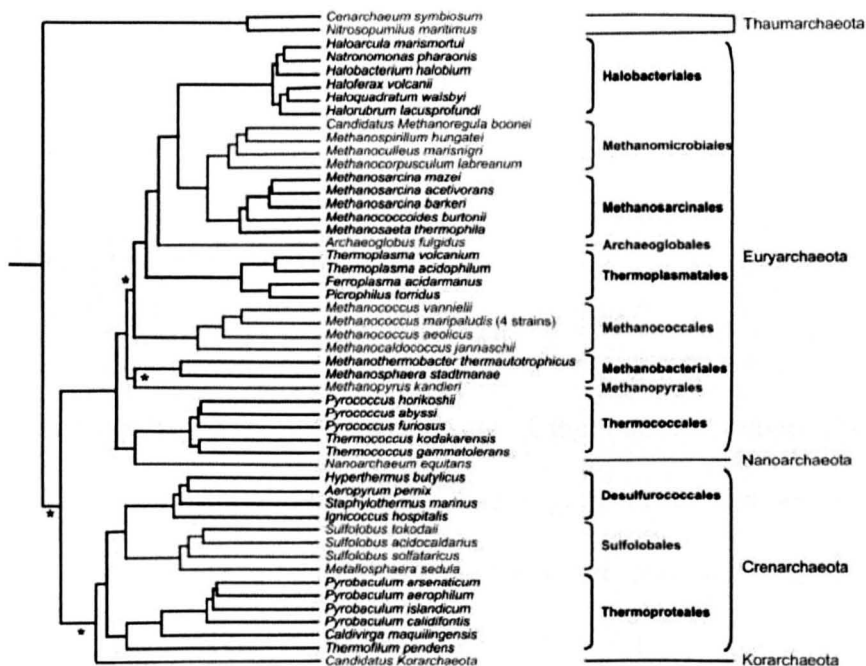
These differing viewpoints have led to numerous representations of the Archaeal phylogeny, including those which do not recognise Nanoarchaeota, Korarchaeota or either as distinct phyla (Fig 1.3).

## 1.4 Crenarchaeota

Crenarchaeota are the most abundant of the archaea, likely having the highest biomass, as well as being ubiquitous in the environment (Hu *et al.* 2011). The crenarchaeota contain many extremophiles, with affinity for both acidity and temperature (Blochl *et al.* 1997). Crenarchaeota display a wide range of lifestyles, including extremes of temperature up to 113°C, *Pyrolobus fumarii* and extremes of pH, as low as pH 1, *Sulfolobus* strains. In addition as well as respiratory metabolism, comprising aerobic growth, anaerobic growth and facultative anaerobic growth, this group of archaea are capable of chemorganotrophism and chemolithoautotrophism (Blochl *et al.* 1997). In addition crenarchaeota appear to have a significant role in the carbon cycle, being a major, numerically predominant component of marine microflora (Hu *et al.* 2011)



A



B

Fig 1.3 A Phylogenetic tree of domain Archaea based on RRS sequences B  
Phylogenetic tree of domain archaea, based on comparison of core genomic  
subsets including horizontal gene transfer events (Gibraldo & Brochier 2009)

## 1.5 Nanoarchaeota

Currently only a single nanoarchaeon is known, *Nanoarchaeum equitans*, which has a total cell diameter of only 400 nm (Huber *et al.* 2002). The organism is symbiotic and grows on another archaeon *Ignicoccus* (Jahn *et al.* 2008). Both the symbiote and its host are hyperthermophilic (Huber *et al.* 2002). Similarities exist between the genome of *N. equitans* and the euryarchaeon *Thermococcus* (Waters *et al.* 2003). Additionally it has the smallest known archaeal genome, coding for genes that affect DNA repairs and recombination, but has none for lipid, nucleotide or amino acid synthesis (Waters *et al.* 2003).

## 1.6 Korarchaeota

Korarchaeota branch close to the root of the archaeal domain (Sarubhai *et al.* 2007). Not much is known about this phylum of archaea, as, at present only the nucleic acids of this group have been detected, culturing them so far has proven impossible (Baker & Cowan 2004). The data available shows the korarchaeota to be hyperthermophiles (Auchtung *et al.* 2006), a conclusion reached by the presence of their DNA in the Obsidian hot spring at Yellowstone National Park (Barns *et al.* 1994).

## 1.7 Euryarchaeota

Euryarchaeota exhibit a large range of lifestyles, with around 190 species, split into 8 orders (Blochl *et al.* 1997). The largest subdivisions of the euryarchaeota consist of methanogenic or halophilic organisms, which are closely related to one another (Blochl *et al.* 1997) (Fig 1.4). Methanogens are characterised by their production of methane gas and have been isolated from a range of environments, including peat bogs, salt marshes and trees (Chaban *et al.* 2006). Regardless of where they have been isolated from they exhibit an anaerobic lifestyle (Chaban *et al.* 2006). Halobacteriales, by contrast are mostly aerobic, with some anaerobes and facultative anaerobes and live in extreme salt concentrations, such as the Dead Sea (Wheelis *et al.* 1992). Interest in these organisms extends to xenobiology, as the extreme nature of their environment is similar to that found on Mars (Landis 2001). Other members of the euryarchaeota include the hyperthermophilic genera *Thermoplasmales* and *Thermococcales* as well as organisms capable of sulphur reduction (Reed *et al.* 2006).

## 1.8 Archaeal Physiology

Archaea have a wide range of morphologies, including several pleomorphic

species (Huber *et al.* 2002). This range, contributed to by the unique nature of archaeal cell walls, includes disks, squares, cocci and bacilli as well as several other shapes (Beveridge 2002). Cell diameters and lengths are also highly varied ranging from 0.4 to 15 µm in diameter, with lengths reaching 200 µm (Huber *et al.* 2002). Additionally they multiply, not only by binary fission as do bacteria, but also by budding, similar to the mechanism used by yeast (Beveridge 2002). Much like bacteria some archaea form aggregates or filaments rather than living as single cells (Desmond *et al.* 2007). The only archaeon implicated in human disease a member of the *Methanobrevibacter* genus is one such, forming colonies in the mouth (Lepp *et al.* 2004).

## **1.9 Archaeal Habitats**

Archaea can be found in an extremely diverse range of habitats (Chaban *et al.* 2006). These habitats range in temperature from 1°C with the organism *Methanogenium firgidum* (Saunders *et al.* 2003) to 121°C with the organism *Pyrodictium* strain 121 (Kashefi & Lovley 2003); in acidity from 0 pH with organisms of the genus *Picrophilus* to 11.7 pH with *Halalkalicoccus jeotgali*

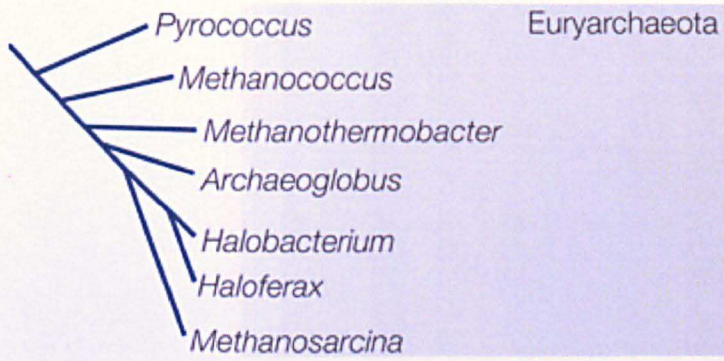


Fig 1.4 Phylogenetic tree of euryarchaeota (adapted from Allers & Mevarech 2005)

(Valenzuela-Encinas *et al.* 2008) and in salinity up to 5.5 M in the order *Halobacteriales*. It has also been shown that archaea inhabit mesophilic environments, such as the human mouth (Lepp *et al.* 2004), soils, trees and fresh water lakes (Samuel *et al.* 2007). Archaea form a significant portion of the global biomass (Olsen 1994 & Karner *et al.* 2001), contributing to the carbon and nitrogen cycles as well as occupying niches unavailable to organisms from the other domains.

### **1.10 Halophilic archaea**

Perhaps one of the most interesting environments inhabited by archaea are the extremely saline seas and lakes. These areas vary in salinity, as high as 5.5M salt (Valenzuela-Encinas *et al.* 2008) and some archaea living in them require salinities as high as 2M for survival (Oren 2002). By contrast mesophilic organisms can rarely tolerate salt concentrations much in excess of 0.2M (Brock 2008). Although there are several none archaeal halophiles these organisms use a compatible solutes system to balance the strong osmotic effects produced by hyper saline environments (Oren *et al.* 2005). In contrast halophilic archaea eschew a compatible solutes system in favour of an iso-osmotic balance with their environment (Oren *et al.* 2005). Whilst this allows them to live at greater salinities than other organisms the high intracellular



levels of ions offer unique challenges to the organism's biology. Biochemical processes including nucleic acid and protein synthesis must be carried out in salinities greater than 2M and as such the systems within halophilic archaea differ from those of other organisms to compensate.

### 1.11 *Haloferax volcanii*

*H. volcanii* (Fig 1.5) is a haloarchaeon belonging to the family *Halobacteriaceae* (Blochl *et al.* 1997). It was first isolated from the dead sea in 1975 (Mullakhanbhai & Larsen 1975) and has optimum growth conditions of 2.6 M salt, 45°C and pH 8.5 (Oren, 2002 & Mullakhanbhai & Larsen 1975).

*H. volcanii* is an obligate halophile, capable of growth only in conditions above 1 M salt, below which concentrations osmotic shock causes cell lysis (Oren 2002). *H. volcanii* is capable of growth at a wide range of temperatures, down to 30°C and above 50°C (Wilson *et al.* 1999). It is a chemoorganotroph and an obligate aerobe. *H. volcanii* has a 42 Mb genome, consisting of a single chromosome and four megaplasms (pHV1, 2, 3 and 4) (Allers 2004).

*H. volcanii*, unlike many micro-organisms is also strongly pigmented, ranging from a pale pink to a deep red in colour. This colouration is caused by its

production of carotenoids, particularly lycopene (Ronnekleiv & Liaaen Jensen 1995). It has been suggested that these carotenoids may provide protection from the high levels of incident ultra violet (UV) radiation found in the dead sea (Dundas & Larsen 1963), however little research has been done into the true role of these compounds in this organism.



Fig 1.5 *Haloferax volcanii* (Allers 2004)

## **1.12 Carotenoids**

Carotenoids are organic pigments, some, such as lycopene and phytoene are hydrocarbons, whilst others such as bacteriorubern and zeaxanthin contain hydroxyl groups or oxygen (Sandmann 2009). Carotenoids have long chain lengths, usually over thirty carbons and are typically very hydrophobic, often with log octanol water partition coefficients of -10 or greater (Britton *et al.* 1998). Carotenoids present a unique spectral signature when analysed with absorbance spectrophotometry. The spectra produced by carotenoids have a three peak structure (Fig 1.7), with distinct maxima for each carotenoid in a given solvent system (Britton *et al.* 1998). These signatures therefore allow identification of carotenoids by comparison to standards, or previously recorded spectra.

## **1.13 Carotenoids in animals**

Carotenoids perform a variety of functions in living organisms. They are utilised and synthesised by all domains of life, although their synthesis is uncommon in animals, with only a single species of aphid, known to produce them (Moran & Jarvik 2010). In animals carotenoids perform a wide range of

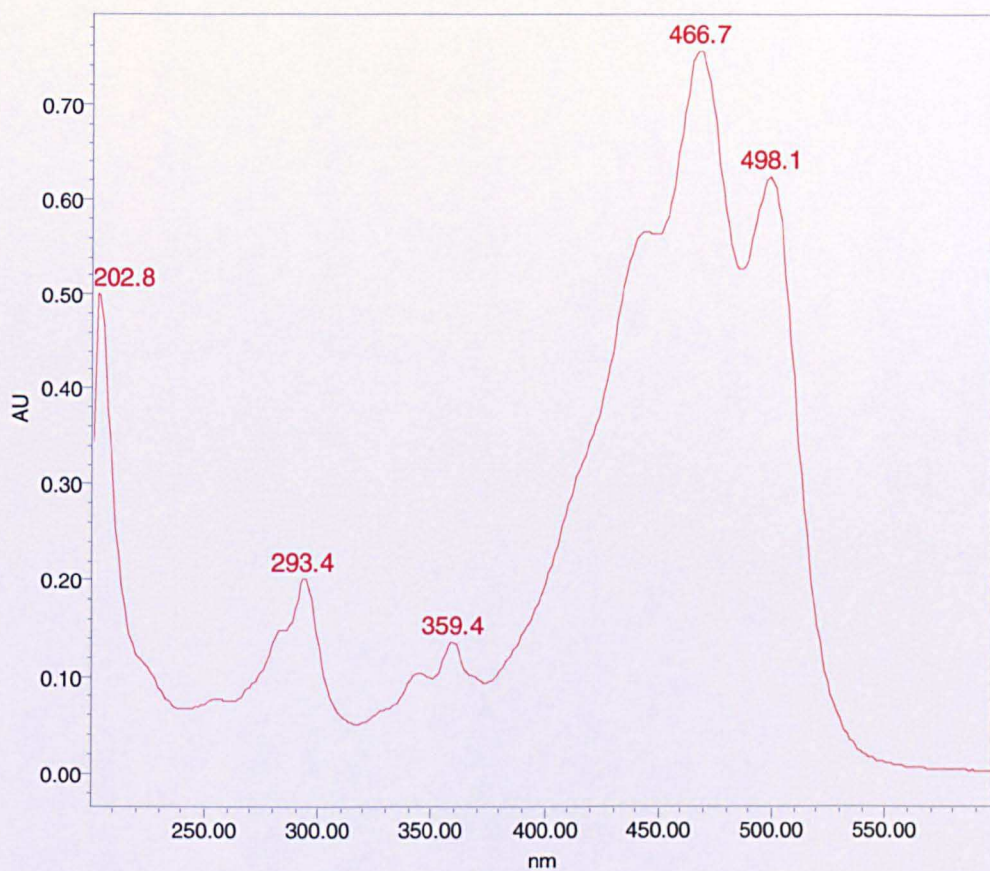


Fig 1.7 Spectrum of lycopene, with maxima at 447.6, 466.7 and 498.1 nm

functions, providing pigmentation in the skin, acting as antioxidants to soak reactive oxygen species (ROS) and as precursors to essential biological molecules, such as retinal, used for sight and retinoic acid, a growth hormone.

### **1.14 Carotenoids in plants**

In plants carotenoids are also used as hormones, with several products of the carotenoid pathway performing this function (Hubbard *et al.* 2010). Additionally they are used to provide colour to various fruits, such as tomatoes. Arguably their most important role, however, is in the chloroplast, where they are involved in photosynthesis. In photosynthesis carotenoids provide two major functions, the first and most studied, is extension of the range of wavelengths at which light can be captured and utilised by the photosynthetic apparatus (Cogdell 1978). The energy absorbed by the carotenoids is transferred to the photosynthetic apparatus via a charged group on a protein (Mimuro & Katoh 1991). The second function performed by the carotenoids in photosynthesis is to act as protective agent against oxidative attack upon the rest of the photosynthetic apparatus (Cogdell 1978).

### **1.15 Carotenoids in bacteria and protists**

In cyanobacteria and photosynthetic algae carotenoids perform the same function as in plants, as part of the photosynthetic apparatus. In addition some bacteria utilise carotenoids to protect against photo-oxidation events both within the cell and exterior to it (Mathews & Krinsky 1965), additionally they are known to provide some protection from incident radiation of the frequencies at which they absorb. It has been shown (Mathews & Krinsky 1965) that bacteria using carotenoids in this way do not produce them when grown in the dark, presumably due to the high metabolic cost (Michal, 1999) being detrimental when no light is present to act on the cell directly or cause the production of ROS. By contrast it has recently been shown that *Deinococcus radiodurans* produces lycopene, a carotenoid, in all growth conditions and utilises it to protect against oxidative attacks from a range of sources, including  $\gamma$ -radiation emitted from nuclear reactors, the presence of the carotenoid greatly bolstering the organisms ability to survive in this extreme environment (Tian *et al.* 2007).

## **1.16 Carotenoids in archaea**

Archaea, too produce carotenoids, particularly among the halophiles. *Halobacterium salinarium* is well known for its purple colouration and ability to perform none photosynthetic photophosphorylation (Danon & Stoeken

1974). Both of these features are a consequence of the presence of bacteriorhodopsin in the cell, a pigmented compound incorporating carotenoids. As well as producing carotenoids for use in photophosphorylation, thus mimicking plants, several archaea synthesise carotenoids for which, at present the purpose is unknown. *H. volcanii* is one such organism, producing lycopene and bacterioruberin in large amounts, with no apparent purpose beyond their accumulation in the membrane (Oren *et al.* 2005).

## **1.17 Carotenoids and Reactive Oxygen Species**

Reactive oxygen species are damaging to organisms as they can react with many important biological molecules, including DNA and proteins, perturbing their function (Paez *et al.* 2011). These compounds can be generated in a number of ways, both metabolically and environmentally (Liu *et al.* 2002, Pedone *et al.* 2004, Kim *et al.* 2011). Through various mechanisms, including auto-oxidation (Pedone *et al.* 2004) and reactions within the electron transport chain of mitochondria and bacteria (Liu *et al.* 2002) ROS are generated metabolically within cells. Additionally they can be generated environmentally by gamma radiation (Kim *et al.* 2011) which acts directly on oxygen molecules, ultra violet radiation (He & Hader 2002) which increases



intracellular ROS production in a range of organisms and Fenton like reactions (Brewer 2010) which produce ROS as a consequence of transition metal chemistry.

Carotenoids including lycopene have been shown to have a protective effect against oxidative damage (Heber & Lu 2001). Their primary role in this regard appears to be the quenching of singlet oxygen species, such as those produced by the mechanisms described above. This quenching ability is of varying strength among the carotenoids, with lycopene being the strongest quencher, although mixtures of carotenoids, particularly those containing either lutein or lycopene, are stronger quenchers than any single carotenoid (Heber & Lu 2001).

The mechanisms of action for the quenching of ROS in carotenoids are varied (Galano & Francisco-Marquez 2009). These reactions can be both degradative, whereby the carotenoid is attacked by the ROS, typically at one of the carbons bordering a double bond, or catalytic in which the carotenoid is not degraded. The catalytic type of activity can occur through electron transport action or hydrogen ion transfer (Galano & Francisco-Marquez 2009). Both mechanisms aid in protecting a cell from oxidative damage.

## 1.18 Carotenogenesis across the Domains

Carotenogenesis across all domains of life shares several characteristics. It starts with phytoene synthesis from geranyl geranyl pyrophosphate (GGPP), which is a product of the isoprenoid pathway. This step is carried out by the highly conserved phytoene synthase (Sieiro *et al.* 2003). Both before and after this step, however, there are differences in the pathways used, particularly between bacteria and eukaryotes with archaea using a wider variety of systems (Sieiro *et al.* 2003; Klassen 2010).

In eukaryotes, and many archaea, including the *Halobacteriales* members that have had the pathway elucidated, the first step in the pathway begins at acetyl-coA (Michal, 1999). This then leads through a number of compounds to isopentyl pyrophosphate (IPP), the first molecule in the isoprenoid pathway. In most bacteria rather than beginning at acetyl-coA the pathway draws from pyruvate and is slightly shorter, again producing IPP (Sandmann 2008). GGPP is then formed from IPP by the highly conserved GGPP synthase (Wiemer *et al.* 2009).

After phytoene synthesis eukaryotes typically use a two enzyme system composed of a phytoene desaturase and a  $\zeta$ -carotene desaturase to produce

lycopene by desaturation of phytoene to  $\zeta$ -carotene, followed by desaturation of  $\zeta$ -carotene to lycopene. In contrast bacteria and many archaea have a single phytoene desaturase (*crtI*), which desaturates phytoene to lycopene in a four step process. In addition to using different enzymes, different intermediates are produced by the two pathways, with the bacterial pathway producing phytofluene, 7,8,11,12 tetrahydralycopene and neurosporene and the eukaryotic system producing  $\zeta$ -carotene and neurosporene.

Despite these differences there are strong similarities between all phytoene desaturases across all three domains (Klassen 2010). This demonstrates that carotenogenesis is an evolutionarily deep branching process, most likely being present in the common ancestor of all domains, with clear evolutionary links between the genes involved. In the case of the two enzyme system in eukaryotes the desaturases (CrtP and Q) are more closely related to each other than any other carotenogenic enzymes (Klassen 2010).

## **1.19 Archaeal Carotenogenesis**

Carotenogenesis in archaea is, at present, little studied. In spite of this it is possible that study of this process could lead to understanding the origin of carotenogenesis across the domains (Sieiro *et al.* 2003; Klassen 2010). As with

all organisms incapable of photosynthesis archaea use the single desaturase model of lycopene synthesis (Yatsunami *et al.* 2004). The most studied role of carotenoids and their synthesis in the archaea, particularly the haloarchaea is related to the synthesis of bacteriorhodopsin (Klassen 2010).

Whilst the early stages of the carotenogenic pathway in archaea mirror bacteria, later steps, particularly those involving cyclases more closely mirror fungi (Klassen 2010). Several archaea contain homologs of the gene *CrtY*, which codes for a carotenoid cyclase more commonly found in eukaryotes (Klassen 2010). These similarities may arise due to the fact that both archaea and fungi appear to have evolved carotenogenesis driven by rhodopsin production.

Some archaea, such as *H. volcanii*, that do not produce rhodopsins do synthesise carotenoids (Sieiro *et al.* 2003). The purpose that these carotenoids serve in such organisms is poorly understood, furthering this understanding is the aim of this thesis.

## **1.20 Potential Applications of This Research**

The research presented here has a number of potential applications beyond the gain of knowledge alone. One such is the possibility of using an understanding

of *Haloferax volcanii* carotenogenesis to design a system by which lycopene and other carotenoids could be produced more cheaply than through the use of tomatoes, the current organism of choice. This would be beneficial for a number of reasons. Lycopene is a powerful anti oxidant and is therefore used as a food supplement in a wide range of products. Additionally it is difficult to purify from tomatoes and therefore pure samples for scientific purposes is expensive, easier production through the use of either *H. volcanii*, a modified *E. coli* or other suitable organism could lead to cheaper lycopene as well as being a potential source of significant financial gain for its producer.

As well as the potential commercial benefits understanding the role carotenoids play within *H. volcanii* could lead to a broader understanding of their role in microorganisms in general. This could potentially lead to novel uses or to a greater understanding of interactions between carotenoid containing organisms and various antimicrobials used to kill them.

## **Chapter 2**

### **Materials and Methods**

## **Chapter 2 Materials and Methods**

### **2.1 Archaeal Strains**

Several *Haloferax volcanii* strains derived from parent strain DS70 were used. DS70 was initially isolated from the dead sea (Wendoloski *et al.* 2001) and cured of the pHV2 plasmid. The derived strains used were *H. volcanii* H26, which has had the PyrE2 gene deleted and *H. volcanii* H53 which has had the PyrE2 and TrpA genes deleted.

### **2.2 Bacterial Strains**

Three *E. coli* strains were used. *E. coli* Top 10 cells (Invitrogen, Paisley UK) were utilised for transformation of ligated products and plasmid propagation, due to their reported high transformation efficiency and plasmid recovery (Invitrogen, Paisley UK). *E. coli* BL21 (DE3) (Invitrogen, Paisley UK) were used as an expression host, as were *E. coli* Rossetta (Invitrogen, Paisley UK). Both these plasmids contain the DE3 prophage allowing expression of the T7 polymerase for use in conjunction with the PET expression system.

Table 1. E. coli genotypes used in this study

Strain	Genotype	Source
Top 10	F <sup>-</sup> mcrA $\Delta$ (mrr-hsdRMS-mcrBC) $\phi$ 80lacZ $\Delta$ M15 $\Delta$ lacX74 nupG recA1 araD139 $\Delta$ (ara-leu)7697 galE15 galK16 rpsL(Str <sup>R</sup> ) endA1 $\lambda$ <sup>-</sup>	Invitrogen
BL21 (DE3)	F <sup>-</sup> ompT gal dcm lon hsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) $\lambda$ (DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	Invitrogen
Rossetta	F <sup>-</sup> ompT hsdS <sub>B</sub> (R <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) gal dcm $\lambda$ (DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) pLysSRARE (Cam <sup>R</sup> )	Invitrogen



## 2.3 Plasmids and Vectors

### *pET28a*

pET28a, a 5369 bp plasmid containing a kanamycin resistance cassette was used as an expression vector. In addition to the Kan<sup>r</sup> cassette it also contains a lac repressed T7 promoter upstream of a multiple cloning site and a lacI cassette. pET28a allows for the expression of both his tagged and none his tagged proteins, the his tag being available at both the N and C termini.

### *pTA131*

pTA131 is a plasmid developed to facilitate gene knockout in *H. volcanii*. pTA131 contains both an *E. coli* origin of replication and a *H. volcanii* origin, allowing for use in *H. volcanii* and easy handling for genetic manipulation in *E. coli*. In addition it contains an ampicillin resistance cassette for selection in *E. coli* and a pyrE2 gene under the control the ferredoxin promoter to facilitate selection for knockouts in *H. volcanii*. Further it contains a multiple cloning site for the insertion of fragments to be used in the construction of knockout vectors.

### *pTA298*

pTA298 contains the *H. volcanii* TrpA gene and promoter located between two BamHI sites, allowing efficient production of knockout vectors in conjunction

with pTA131. In addition pTA298 contains an ampicillin resistance cassette and an *E. coli* origin of replication.

### ***pDR001-006***

This plasmid series represents the production stages of two knockout vectors, pDR003 and pDR006. All are based on the pTA131 plasmid with the following changes. pDR001 contains a clone of the region of *H. volcanii* genome on the 5' flank of the *crtI* gene, located between KpnI and BamHI sites in the multiple cloning site. pDR002 additionally contains the 3' flank located between BamHI and SpeI. pDR003 then contains the TrpA gene and promoter located between these two flanks. Similarly pDR004-006 were built using the *pmdc* 5' flank, 3' flank and TrpA gene with promoter in the same sequence.

### ***pDR007***

pDR007 is based on the pET28a plasmid, with the *crtI* gene inserted between NdeI and BamHI in the multiple cloning site.

## **2.4 Luria Bertani Media**

Tryptone (Oxoid, Cambridge UK), NaCl (Fischer Scientific, Loughborough UK), and yeast extract (Difco, Lawrence US) were combined with water to a

final concentrations of 10g/l, 10g/l and 5g/l respectively. When agar was required 15g/l bacteriological agar (Difco, Lawrence US) was added. The media was autoclaved (121°C, 103kPa, 15mins) before any necessary antibiotics were added.

## **2.5 Antibiotics**

Antibiotic stocks were prepared at 1000x final concentrations required. The final concentrations were as follows, chloramphenicol (Fisher Scientific, Loughborough UK) 35µg/ml, Kanamycin (Fisher Scientific, Loughborough UK) 50µg/ml, Ampicillin (Fisher Scientific, Loughborough UK) 100µg/ml. All antibiotics were used to select for strains carrying plasmids with appropriate resistance cassettes. Where multiple plasmids were used with differing resistance cassettes multiple antibiotics were added. The concentrations of antibiotic were selected as being at least 10x the minimum inhibitory concentration (MIC) of non-resistant strains, whilst being below the MIC of resistant strains produced with the resistance cassettes used.

## **2.6 30% Salt Water**

240g/l NaCl (Fisher Scientific, Loughborough UK), 30g/l  $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$  (Fisher Scientific, Loughborough UK), 35g/l  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  (Fischer), 7g/l KCl (Fisher Scientific, Loughborough UK), were combined in HPLC grade water (Fisher Scientific, Loughborough UK), there pH corrected to 7.5 by the addition of 1M pH 7.5 Tris HCl (Fisher Scientific, Loughborough UK), before HPLC grade water was added to give the above final concentrations.

## **2.7 10x YPC Stock**

12.5g/l Yeast Extract (Difco, Lawrence US), was combined with 2.5g/l Peptone (Oxoid, Cambridge UK) and 2.5g/l Casamino acids (Fisher Scientific, Loughborough UK) in HPLC grade water before 1M KOH was added to adjust the pH to 7.5. Subsequently HPLC grade water was added to give the correct final concentrations before the mixture was autoclaved (121°C, 103kPa, 15mins).

## **2.8 HvYPC Media**

30% salt water was combined with HPLC grade water and 10x YPC stock to give final concentrations of 18% salt water and 1x YPC and autoclaved

(121°C, 103kPa, 15mins). Subsequently CaCl was added to a final concentration of 3mM. Where agar was required 30% salt water and HPLC grade water were combined with 10g/l agar (Difco) and the mixture microwaved until the agar dissolved, the mixture was then immediately autoclaved (121°C, 103kPa, 15mins) prior to the addition of YPC stock and 0.5 M CaCl.

## **2.9 Hv Na-Lac Media**

30% salt water was combined with HPLC grade water such that the concentration would be 18% in the final media. This was autoclaved (121°C, 103kPa, 15mins). The following components were added in order to the specified final concentrations, Tris·HCL pH 7.5 (26 mM), potassium phosphate buffer pH 7.5 (1 mM), Na-Lactate (Fisher Scientific, Loughborough UK), (0.5%), Hv Min Salts (1.2%), Thiamine (Fisher Scientific, Loughborough UK), and Biotin (Fisher Scientific, Loughborough UK), (0.09%), Uracil (5 µgml<sup>-1</sup>) (Fisher Scientific, Loughborough UK). Thiamine and Biotin solution, phosphate buffer and Hv Min Salts were prepared as per Allers *et al.* 2003.

## **2.10 Growth of *H. volcanii* Strains**

*H. volcanii* strains were grown in a Jeio Tech SI 900 shaking incubator (Medline Scientific, Chalgrove UK) at 45°C, 200 rpm in all medias for a time appropriate to the experiment. Growths were performed with 200 ml of media in 500 ml flasks unless otherwise stated.

## **2.11 Harvesting of *H. volcanii* Cultures**

*H. volcanii* cultures were harvested by centrifugation at 9 kg for 30 minutes in a Beckman J2-21 centrifuge (Beckman Coulter Inc, Brae USA). Pellets were resuspended in 5 mls RO water and transferred to 50 ml falcon tubes (BD Biosciences, Franklin USA) for storage.

## **2.12 Extraction of Lycopene**

*H. volcanii* samples were thawed and sonicated in 45 s bursts with 15 s pauses for 5 min at 5 µm amplitude using a Soniprep 150 (Sanyo, Watford UK) 20µg beta carotene (Fisher Scientific, Loughborough UK) was added as an internal

standard. 6.5 ml propan-2-ol was added and the sample vortexed briefly. 5 ml hexane was added and the sample mixed by inversion 5 to 15 times. The sample was centrifuged at 1500 RCF for 30 min in a Megafuge 40R centrifuge (Thermo Scientific, Basingstoke UK) The sample was decanted into a 50ml separation funnel (Fischer Scientific, Loughborough UK) and the lower, aqueous phase run off to waste. The upper organic phase was run into a pre-weighed bijoux, and allowed to dry under air flow prior to storage at -80°C in the dark.

### **2.13 HPLC Separation of *H. volcanii* Extracts**

Extracts were resuspended in 200µl 1:1 tetrahydrofuran (THF):methanol (MeOH) (Fischer Scientific, Loughborough UK). A 20µl sample was injected into the system. HPLC separation was performed using a Phenomonex 5µM C8 Prodigy column 250 x 4.6mm internal diameter (Phenomonex, Macclesfield UK) protected by a 5µM C8security guard cartridge (Phenomonex, Macclesfield UK).

For isocratic HPLC components were eluted using a Waters 2695 Separation System (Waters, Elstree UK). The mobile phase consisted of 1:19

THF:MeOH. Detection of eluants was performed by UV absorption at 450 nm.

For gradient HPLC components were eluted using a Waters Liquid Chromatography System. The mobile phase consisted of a gradient from 1:5 MeOH:H<sub>2</sub>O to 1:19 THF:MeOH over a 15 minute period. Detection of eluent spectra was performed by UV absorption across a range of wavelengths from 200-600nm using a Waters 996 Photodiode Array (Waters, Elstree UK).

β-carotene and lycopene were identified by comparison to the retention time of internal standards. The mass of lycopene present in the original culture was calculated from a comparison of the peak heights of lycopene and β-carotene (Fig 4.3).

## **2.14 Polymerase Chain Reactions (PCRs)**

All PCRs were carried out using the following method with primer sequences, annealing temperatures, extension times and dimethylsulfoxide (DMSO) concentrations as found in Table 2.2. All components were combined to the following final concentrations 1X Phusion PCR master mix (New England Biolabs, Knowl Piece UK), 500 nM forward and reverse primers, 50 ng/μl *H.*



*volcanii* genomic DNA or 25 ng/μl plasmid. Cycling conditions are found in Table 2.3.

## **2.15 *H. volcanii* Genomic DNA Extraction**

*H. volcanii* H26 was grown overnight in HvYPC broth to >0.6 OD<sub>600nm</sub>. 2 ml samples were pelleted at 8 k g for 5 min. The resulting pellets were resuspended in 200 μl ST buffer (Allers 2003), then lysed by gentle inversion with 200 μl lysis buffer (Allers 2003). The lysate was overlayed with 1 ml ethanol (EtOH) (Fisher Scientific, Loughborough UK) and DNA extracted by spooling. The spool was rinsed in 70% EtOH and allowed to dry. DNA was resuspended in 200 μl TE buffer (Allers 2003).

## **2.16 Digestion of PCR Products and Plasmid Vectors**

PCR products and plasmids were digested using the following method with enzymes described in each experiment, all enzymes are FastDigest enzymes (Fermentas, York UK). The DNA sample was digested in a 1x FastDigest buffer (Fermentas, York UK) for one hour at 37°C.

## **2.17 Ligation of Digested PCR Products and Plasmid Vectors for Cloning**

Digested PCR products were treated with Antarctic phosphatase (New England Biolabs, Knowl Piece UK) at 30°C for one hour before being heated to 90°C in order to inactivate the enzyme prior to ligation. Cut vector and plasmid were mixed in 3:1, 1:1 and 1:3 molar ratios of vector to insert in a 1x ligation buffer with T4 DNA ligase (New England Biolabs, Knowl Piece UK) and left overnight on ice.

## **2.18 Electroporation of electrocompetent *E. coli* strains**

25ng of plasmid DNA or 2µl of ligation mixture were mixed with 40µl of electrocompetent *E. coli* cells, chilled in ice (all strains treated identically). This mixture was transferred to a 0.2cm electroporation cuvette, chilled in ice. An electrical pulse (25µF, 2.5kV, 200ohm) was delivered to the mixture by a BioRad Gene Pulser (Bio-Rad, Hemel Hempstead UK). 1ml LB was added at room temperature before the mixture was transferred to a 15ml falcon tube (BD Biosciences, Franklin USA) and incubated for 1 hour shaking at 37 °C 100rpm. Spread plates of 200, 20 and 2 µl were made on LB plates containing

Table 2.2 Primers, DMSO concentrations, annealing temperatures and extension times for PCR.

PCR	Upstream Primer Sequence	Downstream Primer Sequence	Annealing Temperature (°C)	Extension Time (s)	DMSO concentration (%)
<i>crtI</i> ORF	GGTCGATTCGGTccatggATT TATGTTCTCTACG	GAGCGCCggatccGCCGC AC	69	45	20
<i>crtI</i> 5' Flank	GACGAGGAGGTaCCACGC	CTTTAGGCTgGaTCCCTC CTTAC	66.3	25	20
<i>crtI</i> 3' Flank	GTCTCGGAtcCCGACTCGT C	CGCGGAGAAcTaGTGCA TCC	61.8	10	20
<i>pmdc</i> 5' Flank	CGAAATCTTTATCCACGAA TCGTGGTAGTAGCTACC	GTTCGAACGCCCCTAG GACTCGG	61.8	10	10
<i>pmdc</i> 3' Flank	CGGCGGTTCGGATCCGCGC	GCTTCTTCGACGACTGG TACCAGGGCGG	61.8	10	10
<i>crtb2</i> 5' Flank	GCAGCGGTGGGTACCTCT GTCC	GTTTCGTGTGGGATCCC GCGTC	61.8	10	15
<i>crtb2</i> 3' Flank	CGGCGTGATGGATCCGCG AG	CGATTCACTAGTCGGCG TCTGCTCAACG	62.6	10	10

Table 2.3 PCR Cycling Parameters

Step	Cycles	Temperature (°C)	Duration (min)
Denaturation	1	97	2:00
Denaturation	30	97	0:30
Annealing	30	See Table 2.2	0:30
Extension	30	72	See Table 2.2
Extension	1	72	6:00

100 µg /ml ampicillin or 50 µg /ml kanamycin, dependant on the resistance cassette of the transformed DNA. Plates were incubated overnight at 37 °C.

## **2.19 DNA Agarose Gel Electrophoresis**

0.8 to 1.5% (w/v) agarose (Fisher Scientific, Loughborough UK), dependant on the length of DNA to be visualised, was combined with tris acetate EDTA (TAE) buffer (40mM Tris acetate 2mM EDTA) at pH 8.5 and microwaved until all agarose was dissolved. The solution was allowed to cool before ethidium bromide (EtBr) was added to allow visualistaion of DNA. The mixture was then poured into a mould and allowed to set. DNA samples were prepared in loading dye and run at 70-120v in TAE buffer before being visualised by UV transillumination.

## **2.20 Recombinant Protein Expression**

*E. coli* BL21 or Rosetta containing pDR007 was grown to an OD<sub>600nm</sub> 0.4 before Isoproptl-beta-D-thiogalactopyranoside (IPTG) (Fisher Scientific, Loughborough UK) was added to a final comncentration of 0.1 mM in order to induce protein expression. Subsequently a 1 ml sample was taken each hour for four hours. The remaining culture was then harvested by centrifugation at 9kg for 15 minutes.

## 2.21 Poly Arcylamide Gel Electrophoresis

12.5% reducing sodium dodecyl sulphate polyarylamide gel electrophoresis was performed in order to analyse protein expression. A 12.5% resolving gel was made by combining 4.17ml 30% acrylamide (Fisher Scientific, Loughborough UK), 0.5ml 2% bis-acrylamide (Fisher Scientific, Loughborough UK), 3.75ml 1M Tris, 100µl 10% SDS, 1.5ml HPLC grade water, 8.3µl tetramethylethylenediamine (TEMED) (Fisher Scientific, Loughborough UK) and 33µl ammonium sulphate (Fisher Scientific, Loughborough UK). A 5% stacking gel was made by combining 0.85ml 30% acrylamide, 0.35ml 2% bis-acrylamide, 0.63ml 1M Tris, 50µl 10% SDS, 3.15ml HPLC grade water, 10µl tetramethylethylenediamine and 300µl ammonium sulphate. Gels were cast using BioRad Mini Protean III kit (Bio-Rad, Hemel Hempstead UK) as per the manufacturers instructions.

Protein samples were prepared for loading by heating with loading buffer (Fermentas, York UK). Gels were run at 90v in the Mini Protean III apparatus (BioRad, Hemel Hempstead UK). Subsequently they were stained with Coomassie blue (0.25% w/v coomassie blue, 10% v/v acetic acid, equal parts MeOH and water) for 15 minutes before being destained in destain solution (30% v/v methanol, 10%v/v acetic acid) for an hour and visualised on a light box.

## **Chapter 3**

# **Bioinformatic Analysis**

## **Chapter 3 Bioinformatic Analysis**

### **3.1 Introduction**

Understanding how the biosynthetic pathways within an organism progress is very useful when studying the production of compounds by that organism. Bioinformatic techniques allow comparisons to be made between information on the organism being studied and information about other organisms. From such comparisons putative pathways can be built up, which may then act as guides as to what experimentation would glean the most useful information about the organism being studied. In this study such techniques could allow the production of a lycopene biosynthesis model pathway, which would allow more specific questions to be asked about its behaviour.

Such pathways, whilst useful as guides, are limited in scope. Only proteins or genes present in other organisms can be found using this method, if the organism of interest has a novel pathway it cannot be elucidated. Additionally it is possible that some genes identified by bioinformatic comparisons may not be those involved in the pathway of interest in the organism being studied, as they may be related genes with different functions. Furthermore if an organism's entire genome has not been sequenced and published it is possible that all or part of the pathway being investigated will be unavailable to this sort of interrogation.



As well as producing putative pathways specific models can be made of cell components, based upon their sequence. These models, whilst not perfectly accurate, do give some insight as to the nature and possible function of the proteins studied in this way.

The aim of this chapter is to describe the analysis of the *Haloferax volcanii* genome and related proteins with regards to lycopene production through the use of bioinformatic approaches. At the beginning of this study the pathway used by *H. volcanii* to synthesise lycopene was unknown, this chapter shows how a presumptive pathway was constructed through the use of bioinformatic techniques. Additionally this chapter aims to show the analysis of protein models constructed for several enzymes in this pathway.

### **3.2 Carotenoid Biosynthesis Pathway Elucidation**

The biosynthetic pathway by which a compound is produced gives information that can be used to establish its importance to the cell, likely evolutionary origin and even function. For this reason it was decided that elucidating the biosynthetic pathway of lycopene in *H. volcanii* would be a useful step in understanding the production and purpose of lycopene in that organism. To achieve this a bioinformatics approach was used. The genes commonly used in other organisms

were identified from literature (Michal, 1999) and their amino acid and genetic sequences acquired ([expasy.org/sprot/](http://expasy.org/sprot/)). These sequences can be used to identify which of the pathways for lycopene synthesis are utilised by *H. volcanii*.

The *H. volcanii* genome has been sequenced and as such it is possible to perform BLAST interrogations on it with the previously identified sequences. In order to do this each amino acid sequence is entered into a BLAST engine, which compares these sequences to the translated *H. volcanii* genome in order to identify any that match. Once sequences had been identified in this way, they were then used to interrogate a general BLAST engine of non redundant sequences to find if there were proteins they were more similar to than those found in the lycopene synthesis pathway. In both cases BLAST searches were carried out using the blastp algorithm with the BLOSUM62 matrix, gap costs were set at 11, 1. Only genes found to match on both BLAST searches were used to build the pathway. Restricting the genes used in the model in this way helps to remove false positives produced on the first BLAST, which could occur when *H. volcanii* does not contain the gene being searched for. The sequences selected can be found in Table 3.1. and include those from the closely related *Halobacterium salinarium* as well as those from *Lycopersicon esculentum* and *Deinococcus radiodurans*. Where sequences from less related organisms were used the sequences in *H. salinarium* were presumptive or automatically annotated and those from other species more thoroughly characterised.

Table 3.1, Sequences selected for BLAST interrogation.

Sequence Name (Original Organism)	Potential Match in <i>H. volcanii</i>	E Value
Phytoene synthase ( <i>H. salinarium</i> )	CrtB2	1e-113
Phytoene desaturase ( <i>D. radiodurans</i> )	CrtI	1e-47
Phytoene desaturase ( <i>L. esculentum</i> )	None Found	N/A
Carotenoid desaturase ( <i>L. esculentum</i> )	None Found	N/A
Mevalonate pyrophosphate decarboxylase ( <i>H. salinarium</i> )	pMdc	3e-154
Mevalonate phosphate decarboxylase( <i>M. jannaschii</i> )	pMdc	3e-97
Hmg-CoA reductase ( <i>H. salinarium</i> )	HmgR	0.0
Hmg-CoA synthetase ( <i>L. esculentum</i> )	HmgB	2e-93
GGPP synthetase ( <i>D. radiodurans</i> )	IdsA	6e-22
Mevalonate kinase ( <i>H. salinarium</i> )	Mvk	2e-151
Mevalonate phosphokinase ( <i>H. salinarium</i> )	Mvk	2e-132

This series of interrogations showed that *H. volcanii* did not perfectly match any of the previously identified pathways from other organisms, however it was a fairly close match to that found in both *H. salinarium* and *D. radiodurans*. The similarity of this match to an established pathway allows a fair degree of confidence that the model pathway constructed by this method is a reasonable representation of the *H. volcanii* lycopene biosynthesis pathway (Fig 3.1.).

It can be seen that the lycopene synthesis pathway draws from central metabolism via Acetyl-CoA, additionally many of the reactions involved actively use ATP or NADH. The use of these compounds attributed to enzyme reactions is based on reactions studied for homologous enzymes in the literature (Michal 1999), apart from the phosphorylation steps applying either to mevalonate phosphate or isopentyl phosphate. As the enzyme responsible for this step was not identified by the BLAST search this reaction has a single ATP molecule assigned to it as this is required to donate the phosphate used in the reaction.

Prior to the geranyl geranyl pyrophosphate (GGPP) synthesis step the pathway produces isopentyl pyrophosphate (IPP), this compound is known to be the start of the isoprenoid pathway, which leads to the production of many integral cell components, including fatty acids, sterols and compounds used in protein prenylation. It is reasonable therefore to see the pathway to this point as the start of the isoprenoid pathway and the carotenoid pathway as a branch of that larger

schema. As such for the purposes of studying carotenoids it is beneficial to focus on the segment of the pathway after IPP.

All of the BLAST searches carried out to produce the pathway were restricted to matches with an expect value less than 0.1, this value represents the number of sequence hits that could be expected at the quality of the sequence hit in question when searching a database of the size being interrogated. This means that less than 0.1 hits of this quality would be expected by chance taking into account the search parameters (size of database, query sequence, match to target).

All of the sequences selected had an expect value of less than  $10^{-20}$  and as such it is highly unlikely that any of these sequences were unrelated to those used for interrogation.

This segment, comprising CrtB2 and CrtI, processes GGPP into lycopene. The first step is the production of phytoene from GGPP. Four molecules of GGPP go to produce one molecule of phytoene, a carotenoid which absorbs outside of the visible spectrum. Phytoene is then desaturated by *crtI* to produce lycopene. This is similar to the method used by several archaea (Sieiro *et al.* 2003) and bacteria (Grochowski *et al.* 2006) and dissimilar to the two enzyme method of desaturation commonly seen in plants such as tomatoes (Michal, 1999). No

evidence of any carotenogenic enzymes beyond lycopene were discovered by the methods used.

### **3.3 Protein Models**

The lycopene biosynthesis pathway can be seen to draw from central metabolism, starting as it does at Acetyl Co-A. From here it is used as a building block to synthesise mevalonate and thence isopentyl-pyrophosphate (IPP), the first molecule in the isoprene pathway. IPP is formed into geranyl geranyl pyrophosphate (GGPP) and from this point the carotenoid biosynthesis pathway is initiated. The enzymes of the carotenoid synthesis pathway are CrtB2 and CrtI. CrtB2 forms GGPP into phytoene, then *crtI* desaturates phytoene to lycopene, inserting 4 extra double bonds.

In order to gain insight into the function of these enzymes their sequences, translated from the genetic sequences, were modelled onto similar enzymes. I-Tasser uses multiple models based on various proteins with high sequence homology, recursively comparing the models to one another until the best fit is found. The models produced this way (Figs 3.2 & 3.3.) map closely to the enzymes from *Pantoea ananas* (*crtI* Acc P21685) and *Escherichia vulneris* (CrtB Acc P22872) with few areas where a fit could not be achieved. In addition models made using different servers (Lommetts and I-Tasser) map closely to one another. These two facts suggest that the models are reasonably reliable. Upon

examination of the hydrophobicity of the residues in the *crtI* model (Fig 3.2) it can be seen that there is a significant hydrophobic patch in a cleft in the molecule, this suggests itself as the possible active site of the molecule as nowhere else would the highly hydrophobic carotenoid molecules sit. Additionally there are several residues close to this site, which it is possible may be involved in desaturation reactions (244-273), as they match similarly located residues in related proteins that perform those functions (Zhu *et al.* 2005).

Analysis of the alpha helices seen in the *crtI* enzyme indicate the possibility that this enzyme is membrane bound (Table 3.2a.). The scores shown are produced by a combination of several weight matrices, based on statistical analysis of TMbase, a database containing all known membrane spanning protein segments (Hoffman & Stoffel 1993). Table 3.2a suggests that a three transmembrane structure consisting of residues 46-63, 223-245 and 499-517 would give the most likely structure. This would cause the protein to be split into two segments, with a large domain outside and a large domain inside the cell. Such a model accords poorly with that seen in Fig 3.2 and, additionally would disrupt the suggested active site, causing it to be split either side of the membrane. Alternatively a three transmembrane structure consisting of residues 46-63, 469-487 and 499-517 would accord well with the model (Fig 3.2.) and conserve the predicted active site, whilst simultaneously maintaining the preferred orientations of all transmembrane helices involved. Additionally the proximity of 469-487 to 499-

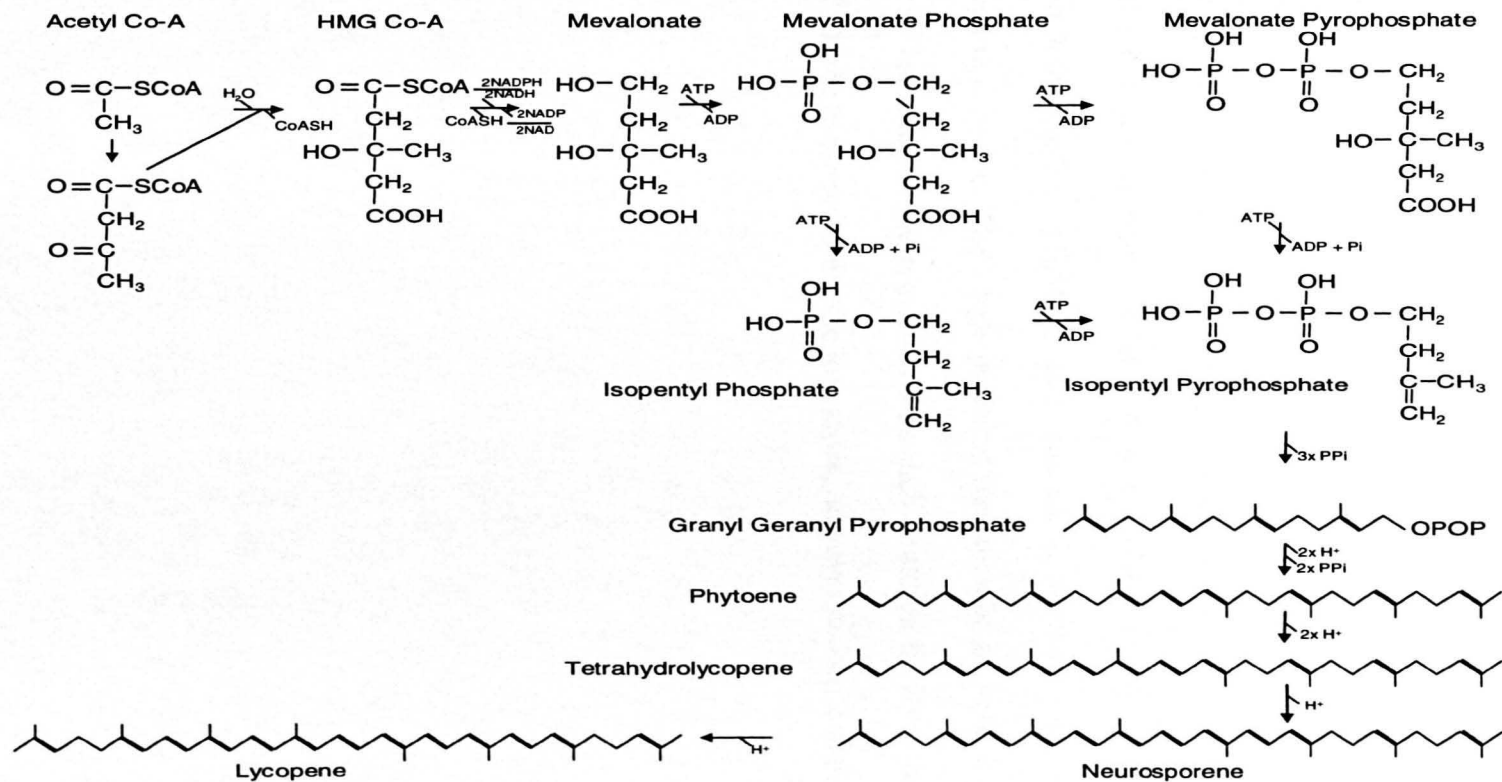


Fig 3.1, The Presumptive Lycopene Biosynthesis Pathway in *H. volcanii*. Labels beside each arrow indicate the enzyme thought to be responsible for the step in the pathway. ATP and NADP use are marked on relevant arrows as is  $\text{H}^+$  and phosphate production.



517 would act to mitigate to some extent the lower score of the former, by ensuring its proximity to the membrane. The final piece of evidence suggesting the second model is more likely is that the helices involved match each other more closely in length, being 18-19 residues long, whilst the 223-245 helix is 23 residues. This analysis whilst showing that the possibility of insertion into the membrane is there, added to the likelihood that this would not disrupt the presumed active site of the protein, does not constitute proof that this is a membrane bound protein. CrtB2 lacks strong evidence of transmembrane helices (Table 3.2b.) making it unlikely that this protein is membrane bound. The single potential transmembrane helix is found close to the C terminus of the protein and as such could be involved in anchoring the enzyme to the membrane, the low score, however makes this unlikely.



Fig 3.2 I-Tasser created model of CrtI, based on a comparison to related proteins showing alpha helices, beta sheets and none ordered regions.

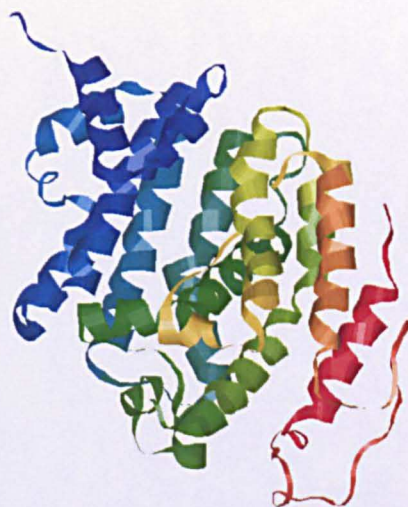


Fig 3.3 I-Tasser created model of CrtB2, based on a comparison to related proteins showing alpha helices, beta sheets and none ordered regions.



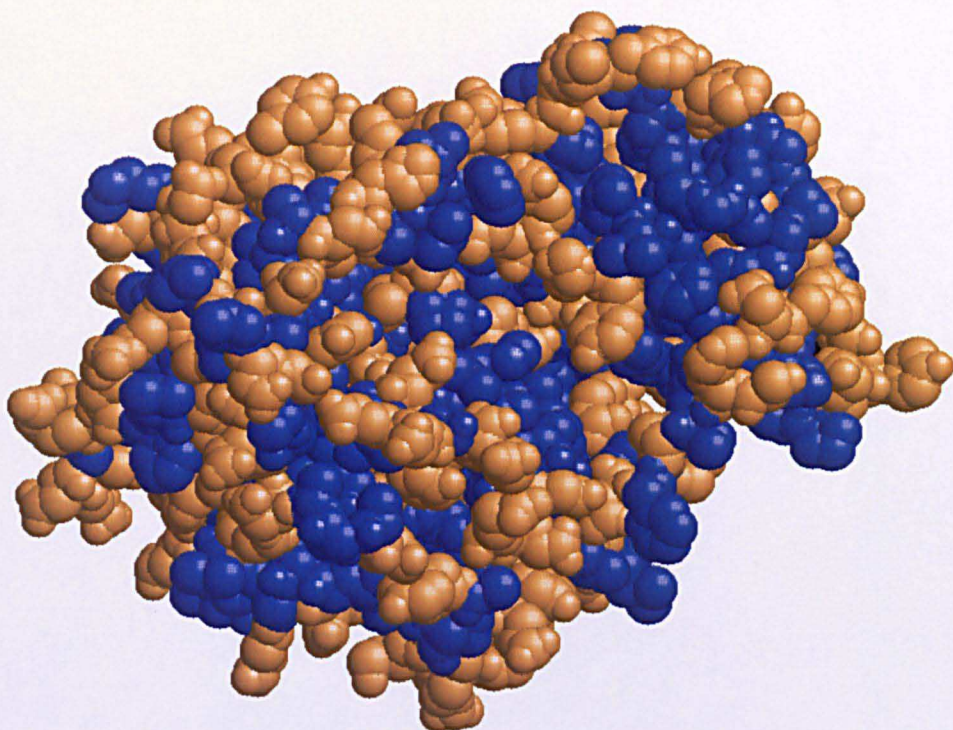


Fig 3.4 I-Tasser created model of CrtI, based on a comparison to related proteins showing a spacefill model with a view of the proposed active site. Blue residues are non-polar, orange polar.

### 3.4 Discussion

From the literature it is clear that there are at least two possible pathways to lycopene from IPP (Sieiro *et al.* 2003), one commonly found in plants and the other more commonly found in bacteria. Additionally we know that there are two paths to IPP from central metabolism, from acetyl-CoA and from pyruvate. The bioinformatic searches performed in this study have indicated, in both cases, which series of enzymes is used by *H. volcanii*. No genes for the pyruvate pathway, nor the carotenogenic pathway featuring multiple desaturases were present and therefore it seems that a pathway starting at acetyl-CoA and proceeding through a single phytoene desaturase is most likely.

The use of expect numbers to restrict matches to those unlikely to be found by chance and the second BLAST of each sequence using the same stringency against the none redundant database increases the likelihood that the sequences found are genuine. Despite this it is not possible to say with certainty that the correct sequences have been identified when using a bioinformatics approach.

Furthering the uncertainty that this is the correct pathway one essential enzyme appears to be missing. The nature of this enzyme, a phosphokinase, is evident, however its substrate is unclear. This is due to the fact that the presumptive mevalonate (pyro)phosphate decarboxylase (*pmdc*) has similarity

when BLASTed to both mevalonate phosphate decarboxylase and mevalonate pyrophosphate decarboxylase. The missing enzyme could therefore either be a mevalonate phosphokinase or an isopentyl phosphokinase. Additionally the mevalonate kinase identified shows good homology to both mevalonate kinase and mevalonate phosphokinase in other organisms, with about equal homology to each, with expect scores in the  $10^{-3}$  range. The evolutionary history of mevalonate kinases contain two distinct families, separately evolved, one from nucleic acid kinases and one from sugar kinases (Smit & Mushegian 2000). Searches carried out with representatives of each family produced no new hits in the *H. volcanii* genome. There is the possibility that Mvk performs both the kinase and phosphokinase role in *H. volcanii*, however study of the protein itself would be required to ascertain if this is true.

PSI-BLAST could have been used at this point in order to improve the accuracy of the results. This could have been useful when assessing the accuracy of the initial BLAST hits in the *H. volcanii* genome. PSI-BLAST replaces the BLAST matrix, in this case BLOSUM with a matrix constructed from an initial BLAST, it then uses this to recursively BLAST and rebuild its matrix until no further hits are found. The output of a PSI-BLAST search can be used to construct a tree of relationships between the proteins found. If this method were used during the second portion of the technique outlined in this chapter in which the *H. volcanii* sequence was re-BLASTed it would have been able to find deeper relationships, however for the purposes of this study the technique used was sufficient.

Each of the steps identified in this pathway has an established cost for the reaction in terms of ATP and NADPH. When these costs are added together they come to 24 ATP and 12 NADPH per lycopene molecule. This cost is high and as such it indicates that it is unlikely that the pathway is a remnant. If it were organisms with mutations in the pathway making it non-functional would be strongly selected for, leading to its loss by the organism.

Additionally it can be seen from the proposed pathway that lycopene biosynthesis runs from phytoene synthesis, subsequent to IPP, thus it is possible that the pathway is regulated at the point of phytoene rather than

Table 3.2a, Transmembrane helices predicted by TMPred for *crtI*. ++ denotes a strong preference for a that orientation, + denotes a slight preference.

Residues	Length	Score	
		Cytoplasm to Cell Exterior	Cell Exterior to Cytoplasm
46-63	18	1071	1384 ++
223-245	23	742	788+
253-282	30	125 +	25
469-487	19	166 +	18
499-517	19	664	962 ++

Table 3.2b, Transmembrane helices predicted by TMPred for CrtB2.

Residues	Length	Score	
		Cytoplasm to Cell Exterior	Cell Exterior to Cytoplasm
249-272	24	84	49



lycopene synthesis, or indeed at both positions. It is unlikely it is regulated at an earlier point in the pathway due to the fact that all other products are used in multiple different pathways and therefore would not be well suited to direct regulation of the carotenoid biosynthesis pathway specifically.

The limitations of this approach, such that only previously characterised enzymes can be used to interrogate the databases, and only those portions of the genome fully sequenced can be interrogated, mean that questions as to enzymes specific to this organism, or of unidentified function cannot be found even if relevant to the pathway in question. Additionally bioinformatic interrogation cannot allow prediction of the function of the products of the pathway, in this case lycopene. Such questions must be answered using other techniques, such as the production of knockouts in the pathway and the cloning and analysis of enzymes. Targets for these techniques are suggested by the results of the bioinformatic interrogation and the presumptive pathway is central to the understanding gained by further experiments.

## **Chapter 4**

# **Lycopene Separation and Quantification**

## Chapter 4 Lycopene Extraction and Quantification

### 4.1 Introduction

In order to analyse any biologically produced compound it is necessary to utilise a method that gives consistent, accurate results. It was decided that due to the small number of carotenoids known to be present in *H. volcanii* (Yachai, 2009) it was appropriate to use high performance liquid chromatography (HPLC) as a method of separation and UV-Vis absorbance as a method of quantification. These combined techniques take advantage of the highly hydrophobic nature of lycopene as well as its unique absorption profile. They do, however require that any other compounds that might mask the lycopene signal are removed from the sample. Fortunately most compounds likely to interfere with analysis of the UV-Vis signal (such as proteins) are hydrophilic and as such can be separated with relative ease from hydrophobic components.

The extraction of lycopene from *H. volcanii* has been performed by several methods, but almost exclusively as part of membrane extraction, rather than for the purpose of studying lycopene itself. Lycopene extraction from other organisms, such as tomatoes, is carried out routinely, however these procedures have been developed to extract carotenoids from plants, which have a markedly different physiology to archaea. For this reason it was

necessary to develop an extraction and quantification method for use in this project in order to produce reliable, replicable data.

## **4.2 Lycopene Extraction**

In order to analyse the levels of lycopene within the *Haloferax volcanii* cell, it is first necessary to disrupt the cell membranes to allow access to the compound. We therefore developed a suitable method of lysis derived from the many different methods that are available in the literature (Dyall-Smith 2009).

The most common method of lysis used with *H. volcanii* when performing molecular biology, is through the use of a lysis buffer consisting chiefly of water and sodium dodecyl sulfate (SDS) as a surfactant (Dyall-Smith 2009). Whilst this method is efficient in lysing the organism, the use of SDS is undesirable as it can contaminate the final sample, potentially altering partitioning properties as well as running properties in HPLC. Fortunately *H. volcanii* can be lysed by exposure to pure water: due to its high intracellular ionic concentration, the introduction to pure water causes lysis by osmotic shock (osmolysis). This lysis, however, may not make lycopene fully available for extraction due to membrane fragments being left intact (Klein *et al.* 2005). It is therefore possible that further treatments may be necessary to give access to the lycopene. Physical methods such as sonication, freeze thaw and glass bead treatment are preferable to chemical methods to achieve this due to the lack of contamination, caused by these techniques.

Once cell disruption has been achieved it is then essential to separate the lycopene from other cell components to allow its quantification and analysis. The most important quality of lycopene with regards to its comparison to other components of a halophilic cell is its high hydrophobicity (Log P = 17.64, Meylan & Howard 1995). The hydrophobic nature of lycopene potentially allows for its extraction by partitioning from the aqueous to the organic phase of a solvent matrix. Access of the solute to the solvent is essential when performing an extraction of this kind. In order to increase this, greater mixing energy and other solvents miscible with both phases can be used to aid this. It is best, however, to use as few solvents as possible both in the interests of simplicity and to reduce possible contamination of the final extract. Hexane, a highly hydrophobic solvent, has been used in extraction of lipids and carotenoids from several sources (Avelano 1988, Sosulski & Gadan 1988), and is therefore a good starting point for designing an extraction system. If a secondary solvent is required to aid miscibility and accessibility, propan-2-ol is an appropriate choice due it being both water miscible and miscible with all organic solvents, including hexane.

A series of extractions were performed on pellets of *H. volcanii*. These pellets were harvested from stationary phase cultures (Section 2.), the stage of growth is when the greatest colouration was evident. These extractions were intended to develop the simplest and most efficient system for extraction of lycopene from *H. volcanii*. The first extraction method called for the mixing of *H.*

*volcanii* lysate, produced by osmolysis, with hexane. Osmolysis was carried out by the resuspension of pellets in 10 ml H<sub>2</sub>O. An equal volume of hexane was added to the lysate and mixed by inversion, 10-20 times, before being allowed to separate (Method A). This method resulted in some separation, but overall had a low efficiency (Table 4.1). The measure of efficiency was made by comparison of two successive extractions of the same sample. An efficient extraction would have the majority of the lycopene in the primary extract whilst an inefficient would have as much in the secondary extract, due to little lycopene being extracted initially thus allowing more to be extracted on a second attempt. The mass of lycopene in the samples was determined by HPLC (Section 2.12). In addition to being inefficient, this extraction method was also very time consuming as separating the aqueous and hexane into distinct layers took a considerable period of settling, typically 24 hours for 100 ml total volume. To improve the efficiency of the extraction, two alterations were made to the method. The first was the addition of 2 lysate volumes of propan-2-ol to the lysate, before the addition of one volume of hexane (Method B). This had several beneficial effects, in addition to improving miscibility; it also caused many hydrophilic cell components to precipitate out of solution, making their separation from the hydrophobic components easier. The second alteration was the addition of a centrifugation step (1500 RCF, 30 minutes, Megafuge 40R, Thermoscientific) after mixing with hexane, rather than allowing the mixture to separate under gravity. These two procedures resulted in an increase in efficiency to 71.3% compared to the 53% of the previous method. Despite these changes there was, however, still a substantial

fraction of lycopene remaining in the secondary extract (Table 4.1). The addition of the centrifugation step and propan-2-ol also led to the development of a new challenge to the extraction procedure, a pellicle now formed at the interphase between the two solvent layers, making it difficult to remove the organic layer without disrupting this body. There was, however, a significant benefit to this technique in that the material precipitated by the propan-2-ol formed a pellet, ensuring it did not interfere with the organic layer. It is likely, therefore, that any lycopene remaining in the aqueous phase would be contained in the pellicle. This most likely arose from the incomplete membrane disruption at the osmolysis step which leaves pieces of membrane intact. To remedy this, a physical method of lysis, freeze-thawing, was added between the osmolysis and solvent addition steps. The samples were frozen to  $-80^{\circ}\text{C}$  for 2 hours and then thawed to allow ice crystal formation to disrupt the cell (Method C). This met with some success as pellicle formation was reduced, although not eliminated, and the efficiency of extraction increased greatly, to 89.5% (Table 4.1). This result informed the decision to add in a second physical lysis step, this time between the osmolysis and freezing steps: that of sonication (Method D). Sonication was performed in 45 second bursts with 15 minute pauses for 5 minutes at 5  $\mu\text{m}$  amplitude (Soniprep 150, Sanyo). Whilst this primarily disrupted cell membrane sections it had the additional benefit of fragmenting genomic deoxyribose nucleic acid (DNA), which significantly reduces lysate viscosity. The addition of the sonication step did, as anticipated, increase the efficiency of the extraction even further (Table 4.1).

Whilst extraction levels and efficiency with Method D were high, it was considered that it might be possible to improve them further by optimising the miscibility of the solvents. This could be accomplished through the varying of the propan-2-ol proportion. A range of propan-2-ol ratios were tested from 1.4 to 2.4 lysate volumes, and a 1.6 lysate volume was found to be the optimum for efficiency (Fig 4.1). Propan-2-ol proportions of less than 1.6 may be insufficient to allow the full access of the hexane. Whilst those greater than 1.6 may allow lycopene to stabilise in the aqueous phase or become associated with cell components that are then pelleted.

Once an optimal propan-2-ol contribution was determined it was considered that a smaller hexane volume would be beneficial for ease of handling and therefore a further experiment was carried out, halving the volume of hexane used in the process. This showed very similar efficiencies to a full lysate volume of hexane, with 71  $\mu\text{g}$  being recovered in the primary extract, as compared to the 73  $\mu\text{g}$  seen with 1 lysate volume of hexane and both showing less than 1.5% of this in the secondary extract. As such the lysate:propan-2-ol:hexane ratio of 1:1.6:0.5 was selected for all future work for both high efficiency of extraction and ease of handling.



Table 4.1, Extraction Efficiencies by Method. Yields (in  $\mu\text{g}$  per 250ml stationary phase culture OD600nm 2.0 52 hours after inoculation) given by different extraction procedures, showing yield given when performed directly on lysate, primary extract and when performed on aqueous phase left over by extraction procedure, secondary extract. Extraction efficiency calculated as percentage of total extract in the primary extract. Standard errors are calculated from 4 replicates.

Technique	Lysis Procedure	Solvent Ratio Lysate:Propan-2-ol: Hexane	Primary Extraction Yield ( $\mu\text{g}$ )	Secondary Extraction Yield ( $\mu\text{g}$ )	Extraction Efficiency (%)
A	Water	1.0:0.0:1.0	11.4 $\pm$ 2.1	10.1 $\pm$ 2.0	53.0
B	Water	1.0:2.0:1.0	29.4 $\pm$ 5.1	10.9 $\pm$ 2.3	71.3
C	Water, Freezing	1.0:2.0:1.0	43.3 $\pm$ 4.7	5.1 $\pm$ 2.2	89.5
D	Water, Sonication, Freezing	1.0:2.0:1.0	69.7 $\pm$ 1.8	4.2 $\pm$ 0.5	94.3

### 4.3 Lycopene Separation and Quantification

Although lycopene extraction removes the majority of cellular components, lycopene is still not in a pure state. In addition, in order to measure the quantity of lycopene in any given sample it is necessary to be able to separate it from these cellular components. There are several methods available for the separation of similar compounds based on a variety of characteristics, the most common of these separate compounds based on their mobility in various solvent systems. High performance liquid chromatography (HPLC) and thin layer chromatography (TLC) are the most common methods used with hydrocarbons, such as carotenoids (Kachick *et al.* 1986). TLC allows separation of compounds according to their mobility in multiple solvent systems, giving an array of compounds separated along two axes. HPLC similarly separates compounds by using their mobility in varying solvents. HPLC takes advantage of a column based approach to separation. This approach is useful in that it allows variation of separation parameters based on a stationary phase and a mobile phase, which leads to great selectivity. Normal phase HPLC separates compounds based on their affinity for a non polar mobile phase over a polar stationary phase, this causes elution in order of polarity, with the most polar eluting last. Carotenoids

are highly non polar compounds. RP-HPLC is commonly used to elute hydrophobic compounds, as with a non polar stationary phase and polar mobile phase polar compounds are eluted first, removing them from the system. RP-HPLC also gives better resolution of non polar compounds (Bakalyar *et al.* 1977). Another advantage of HPLC over TLC is that it allows a more quantitative approach to be taken, as each compound is able to be passed through a detecting apparatus after separation. The choice of detection apparatus is important in determining accurate identity and quantity of compounds. In the case of carotenoids the best choice is a spectrophotometer, as carotenoids have clear and distinct signals in an area of the ultraviolet-visible (UV-VIS) spectrum relatively devoid of interference from other organic components. This unique absorbance pattern is attributed to their large number of conjugated bonds, which absorb light at specific frequencies (Britton *et al.* 1998). In the case of carotenoids the patterns produced are a distinctive three peak structure in the visible spectrum. Visible light is what gives carotenoids the distinctive colours that they are best known for. Each carotenoid has a distinct visible spectrum and as such it is possible to distinguish them based on their absorbance patterns (Britton *et al.* 1998). In order to do this it is necessary to use a photo diode array to scan a range of visible wavelengths as the sample moves past it. This allows a spectrum to be obtained for each peak on the chromatogram, thus

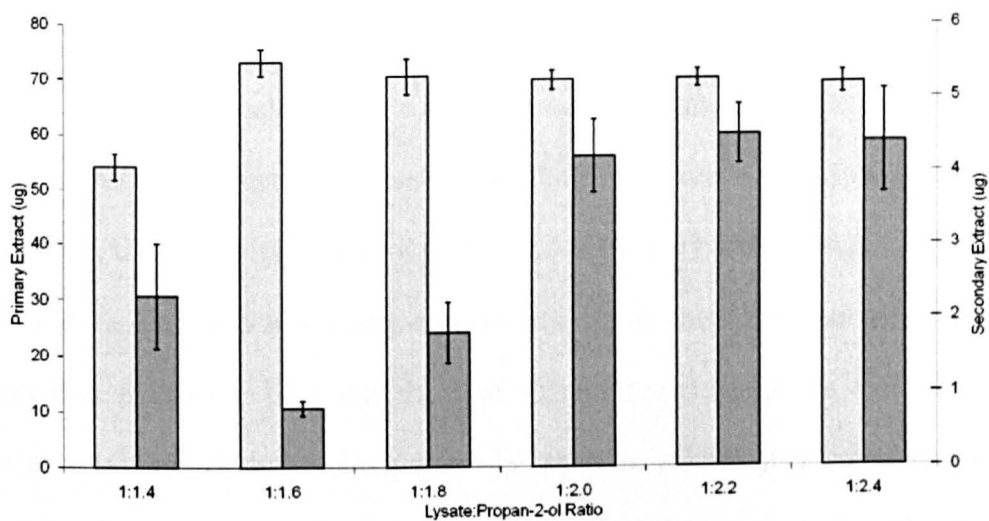


Fig 4.1, Variance in extraction yield with altered lysate:propan-2-ol volume ratios. One volume of hexane was used in all instances. □ Mean primary extraction yield. ■ Mean secondary extraction yield. Yields shown in  $\mu\text{g}$  per 250ml stationary phase culture OD600nm 2.0 52 hours after inoculation, error bars representing standard error calculated from 4 replicates.

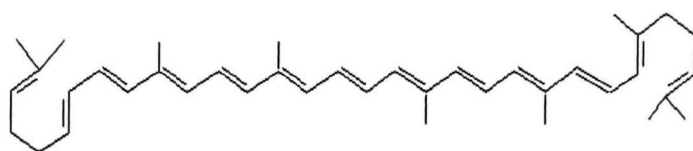
allowing identification of particular carotenoids. In order to separate lycopene from the other extracted components it is first necessary to solubilise the entire final extract in an appropriate solvent. In order to gain an insight into the best solvent to use attempts were made to solubilise  $\beta$ -carotene >95% purity (Sigma-Aldrich, UK), a substance with a similar hydrophobicity and structure to lycopene (Fig 4.2), but much lower cost, in a variety of solvents. Methanol >99% purity (Fischer Scientific, UK) was the first solvent tested as this is low cost and relatively safe. It was found, however, that very little  $\beta$ -carotene could be dissolved in pure methanol and addition of tetrahydrofuran (THF) >99% purity (Fischer Scientific, UK) improved solubility. Using this method it was found that  $\beta$ -carotene could be solubilised at 1 mg/ml in 1:1 THF:methanol. When lycopene was later solubilised in this same 1:1 ratio from both pure stocks and samples it was found to go completely into solution.

After choosing a solvent, column choice is the next essential step. Initially a 5  $\mu$ m C8 prodigy column 250mm x 4.6mm internal diameter (Phenomenex, Macclesfield UK), was selected as it exhibits less retention of hydrophobic molecules than a longer chain (C18) column. A lower amount of retardation is favourable in situations, such as that presented by lycopene, where the compound of interest exhibits significant hydrophobicity compared with the

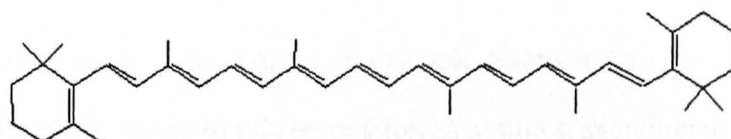
majority of other substances passing through the column.

Methanol was chosen initially as a suitable solvent for the mobile phase however it was found that lycopene failed to elute, likely due to the mobile phase being less favourable hydrophobically than the column. Tetrahydrofuran (THF) was again added at low concentrations until it was found that 1:19 volume:volume ratio THF:Methanol could elute the standards quickly with lycopene eluting at 205 seconds and  $\beta$ -carotene eluting at 260 seconds, cleanly with a symmetrical narrow peak and reproducibly, giving a CV of 1.2% for peak height with 4 replicates from the C8 column.

A standard curve allows a comparison of lycopene to  $\beta$ -carotene peak height to determine total initial lycopene mass in the sample, allowing  $\beta$ -carotene to be used as an internal standard in extractions. The standard curve was produced by dissolving  $\beta$ -carotene and lycopene in 1:1 THF:MeOH to final concentrations of 0.1 mg/ml for  $\beta$ -carotene in all cases and a range of concentrations from 0.025 to 5 mg/ml of lycopene and running these standards using the isocratic HPLC method (Section 2.12). The standard curve also allowed the detection limits of the system to be tested, with the lowest level trialled, containing 0.025 mg/ml of lycopene, still being accurately detectable (Fig 4.3). As the final extracts from



**Lycopene (C<sub>40</sub>H<sub>56</sub>)**



**β-carotene (C<sub>40</sub>H<sub>56</sub>)**

**Fig 4.2, The chemical structure of lycopene and β-carotene**

cultures would be solubilised in 200  $\mu$ l 1:1 THF:MeOH 20  $\mu$ g of  $\beta$ -carotene would be added to each prior to extraction (Section 2.11) After the establishment of the standard curve a culture of *H. volcanii* was grown up in HvYPC overnight (Section 2.9) to allow the system to be tested with a real extract. Lycopene was extracted using the method developed above (Section 2.11) (Fig 4.4). 3 replicates were processed identically and gave similar results, CV 3% on lycopene mass determined to be in the sample, an example trace can be seen in Fig. 4.5. Whilst the method developed to this point allowed accurate quantification of lycopene, in principle, it did not allow identification of the peaks by anything other than retention time. This flaw meant that it was possible that a compound that eluted at the same point as lycopene could cause errors in the results.

For this reason it was decided that a Waters 996 photo diode array detector (PDA) (Waters, Elstre UK) would be used to record the spectrum at each point on the chromatogram over a range of wavelengths. These spectra can then be used to analyse the make up of a sample by comparison with the known spectra for carotenoids. Once the PDA was linked into the system a sample was injected and the subsequent trace produced analysed (Fig 4.5.). Comparison of the spectra of the presumed lycopene to the literature (Britton *et al.* 1999) confirmed that it was indeed lycopene.



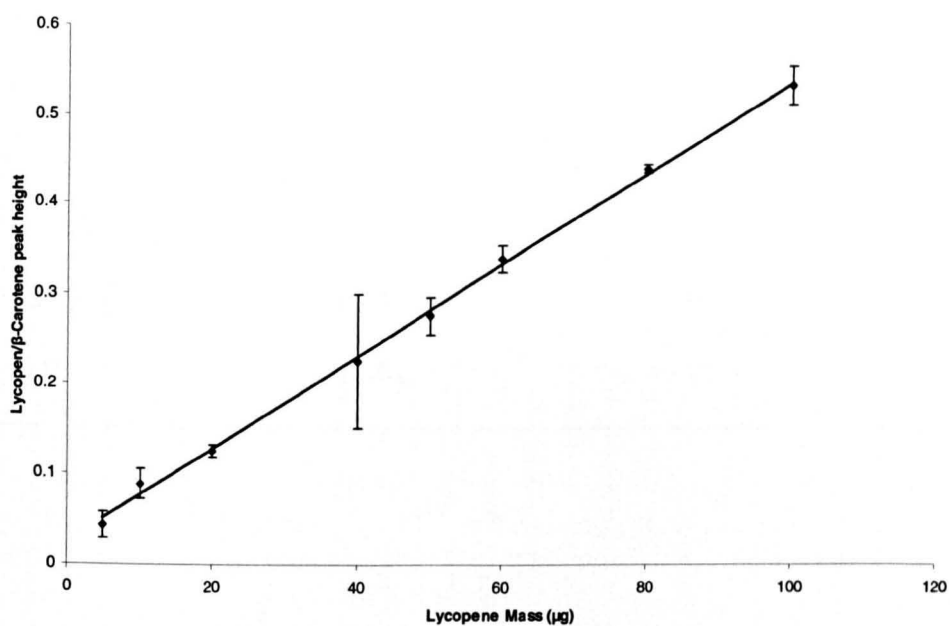


Fig 4.3, Total lycopene mass in the sample compared to lycopene absorbance (450nm) peak height/ $\beta$ -carotene absorbance (450nm) peak height with 0.1 mg/ml  $\beta$ -carotene. Standard error bars are computed from 4 replicates. HPLC apparatus set up as per the isocratic method for lycopene analysis (Section 2.12).  $R^2$  0.99

Table 4.2, Chemical properties of lycopene and  $\beta$ -carotene

Chemical Property	Lycopene	$\beta$ -carotene
Formula	$C_{40}H_{56}$	$C_{40}H_{56}$
Molecular Weight	536.87	536.87
Density (g/cm <sup>3</sup> )	0.941	0.888
Log P	17.64	17.62

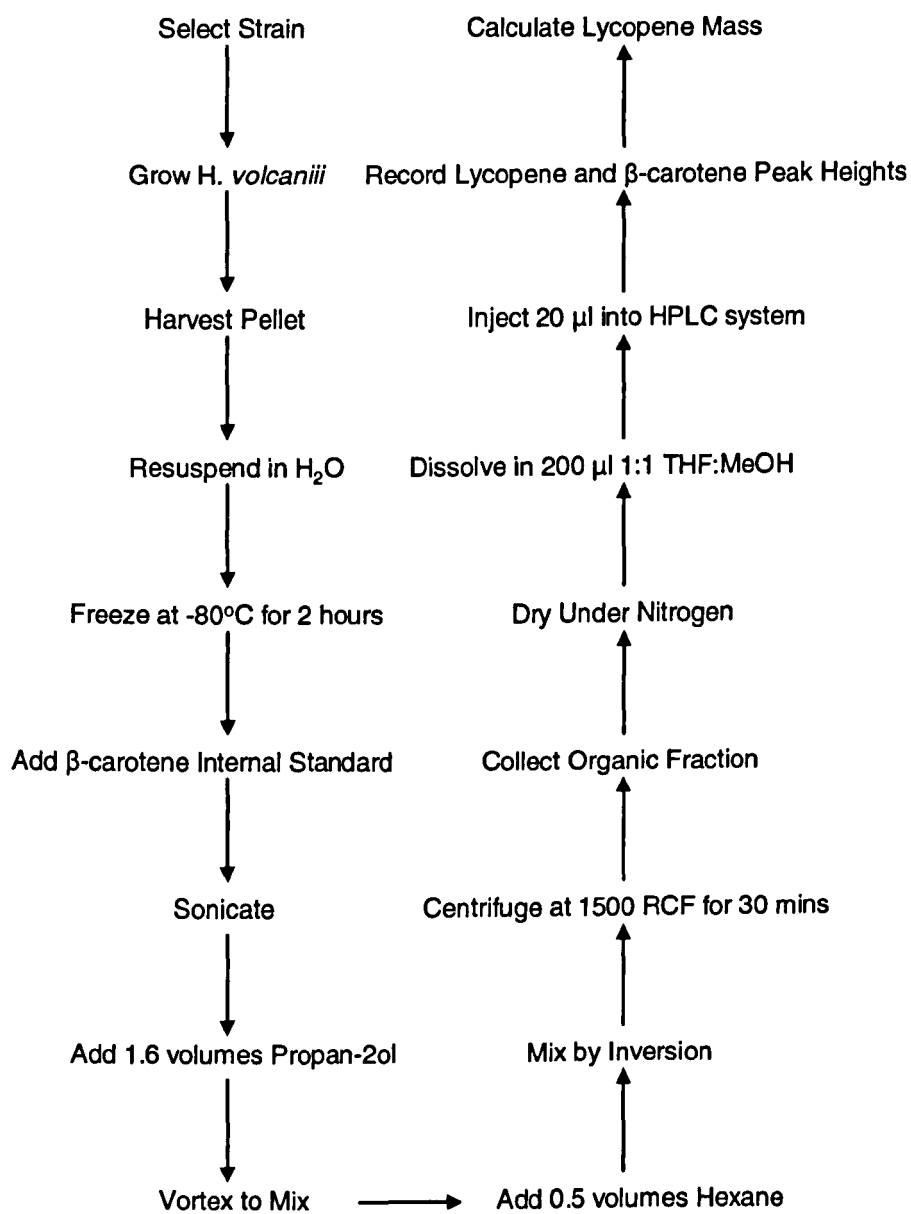


Fig 4.4 Extraction and Quantification of Lycopene from *H. volcanii*

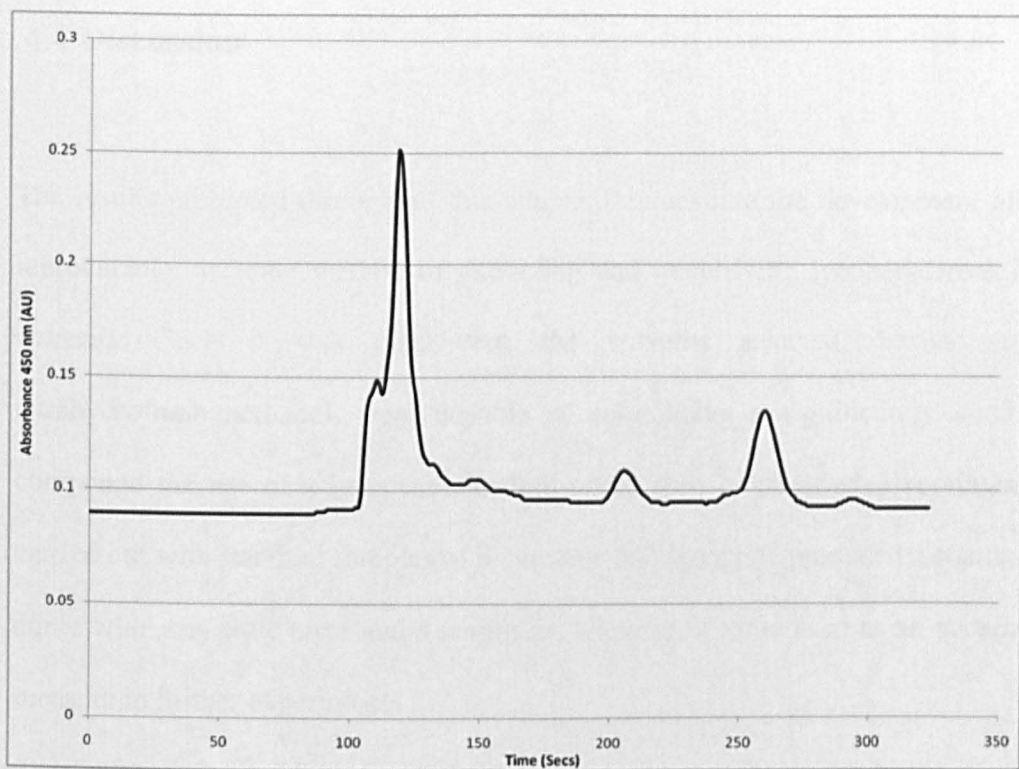


Fig 4.5, A chromatographic trace of a *H. volcanii* extract taken from *H. volcanii* H26 showing peaks corresponding to the retention time of lycopene (205 secs) and  $\beta$ -carotene internal standard (260 secs)

## 4.4 Discussion

The results presented throughout this chapter demonstrate the development of a reproducible, accurate method of extracting and quantifying lycopene from *H. volcanii*. Once it was clear that the solvents selected, hexane and tetrahydrofuran:methanol, were capable of solubilising a significantly similar compound the use of a lycopene standard could then be justified. Experiments carried out with standard samples of  $\beta$ -carotene and lycopene produced a standard curve with very little error and a strong fit, allowing it to be used as an accurate measure in further experiments

During optimisation of the extraction procedure it was noted that propan-2-ol balance was important, if either too much or too little was added this resulted in a reduced extraction efficiency. It is likely that too little propan-2-ol resulted in insufficient mixing between the organic and aqueous phases, leading to reduced solvent access and therefore lower efficiency. The reduced efficiency with higher propan-2-ol levels is unexpected. It is possible that higher proportions of propan-2-ol assist in stabilisation of the lycopene in the aqueous phase, reducing the amount transferring to the recovered organic phase. It is also possible that propan-

2-ol mixed with hexane would cause the organic phase to be more polar, these two factors may reduce the difference in hydrophobicity between the phases, which could contribute to reduced lycopene yield.

In addition to altering the miscibility of the solvents and increasing organic solvent access to the compounds of interest propan-2-ol also caused the formation of a precipitate upon its addition to the aqueous phase. It is reasonable to assume that this precipitate was formed of hydrophilic cell components (particularly DNA and proteins) which are known to precipitate out of solution in the presence of alcohol. The formation of this precipitate can therefore be seen as advantageous due to it reducing the chance of such components being carried over into the final sample and potentially contaminating the results.

The choice of a photodiode array detector for the analysis of carotenoids was based upon a number of factors. The primary alternate method would be liquid chromatography mass spectrometry LC-MS, which, whilst applicable has a higher technical investment and was not as readily available as the PDA for use in this study. LC-MS gives the same values of  $m/z$  for some carotenoids, such as lycopene and beta carotene with the same chemical formula (Hagiwara T *et al.*

1998), whilst their different elution points would still allow determination of which carotenoid was being detected, it would rely on knowledge of retention times, much as the absorbance method without a PDA does. Additionally cholesterol based compounds and other lipids can interfere with LC-MS operation (Hagiwara T *et al.* 1998) and would require a greater time investment in optimising the technique. Despite these flaws an LC-MS method would allow more direct and accurate quantification of extracted carotenoids. It was decided that for the purposes of this work the PDA method was more appropriate as it allowed a higher throughput, due to greater equipment availability, offered the advantage of direct identification of carotenoids based on absorbance and required less optimisation than an LC-MS method.

## **Chapter 5**

# **Changes in the Carotenoid Metabolite Profile of *Haloferax* *volcanii* Throughout Growth and Under Stress Conditions**



## **Chapter 5: Changes in the Carotenoid Metabolite Profile of *Haloferax volcanii* Throughout Growth and Under Stress Conditions**

### **5.1 Introduction**

Micro organisms can be partly characterised by their growth behaviours, doubling time, lag phase and growth rate in various media. In order to study a microbe it is important to ensure that these factors are not only known, but also reproducible across the experiments performed. With this in mind this chapter shows the reproducibility and timings of *H. volcanii* growth in all media used during this project. It also justifies the use of more than one type of medium, due to the need to perform experiments with more carefully controllable culture conditions.

Once culture conditions and media to be used have been established this chapter then goes on to examine the production of lycopene during the various stages of cell growth. From here it proceeds to attempt to elucidate the possible roles of lycopene in the cell. The roles hypothesised and tested are based upon the

chemical and physical properties of lycopene, its extreme hydrophobicity, light absorbing ability and reductive nature. Each possibility is examined and from the data collected a likely role for lycopene, as well as some suggestions on the nature of its regulation are put forward.

In addition to the study of lycopene throughout cell growth and under varied stress conditions the carotenoid profile of the archaeon is also examined, with insights into the carotenoid pathway and its intracellular behaviour gained through this. All cultures in this chapter were grown as per Section 2.9, extractions were performed and lycopene levels measured as per Sections 2.10-2.12

## **5.2 Haloferax Growth Under Non-Stress Conditions**

*H. volcanii*, as most micro organisms, grows in a distinct pattern in liquid culture. This growth is characterised by an initial lag phase, in which little or no growth occurs as the cells acclimatise to their new conditions. This is followed by exponential phase, in which rapid cell growth and division are prominent as the organism makes use of the abundant nutrients available in the media. Next is the stationary phase in which few nutrients are available and the organism enters a

mode of growth best suited to starvation conditions. The final phase is death phase in which the majority of nutrients have been exhausted and the number of organisms decline as they die. As it is in exponential phase that the organism is at its most active metabolically it was decided that this would be the best starting point for investigation.

In order to ensure that experiments conducted during the course of this investigation were in line with others performed upon this organism, it was necessary to establish protocols for handling *H. volcanii*. This was undertaken by analysing the literature characterisation of *H. volcanii* growth in liquid media before selecting a set of conditions appropriate to this investigation. As such growth has been well characterised in HvYPC (Zhou *et al.* 2008) each strain used in the investigation was grown in this broth and its optical density across a time course recorded to establish similarity with published material and reproducibility (Fig 5.1a). All other strains displayed similar growth, although in the case of the mutant strains the already variable lag phase was considerably extended (Table 5.1a). Apart from lag phase length variance all growth curves were highly similar in appearance (Fig 5.1b) regardless of strain and had only trivial differences between replicates. These results imply that working with *H. volcanii* in HvYPC produces reproducible growth within strain and is a reasonable medium to use for

cross strain comparisons. To determine if a culture has entered the exponential phase doubling must be established. To accomplish this it is important to validate any method used by comparison to viable count. In the case of *H. volcanii*, optical density was selected as the most efficient, rapid means of establishing a point in the growth phase. This was compared to viable count measured using the miles misra technique and through direct visualisation under a fluorescence microscope with use of the Bac-Light live dead kit (Invitrogen, Paisley). There is a strong correlation between optical density and viable count measured by either miles misra ( $R^2 = 0.99$ ) (Fig 5.2a) or Bac-Light ( $R^2 = 0.98$ ) (Fig 5.2b). This correlation shows that optical density is a good indicator of viable growth during exponential phase in *H. volcanii* accordingly it was decided that this should be used as the measure by which culture growth phase should be established. With this criterion selected it was then possible to move forward with characterisation of lycopene production in exponential growth.

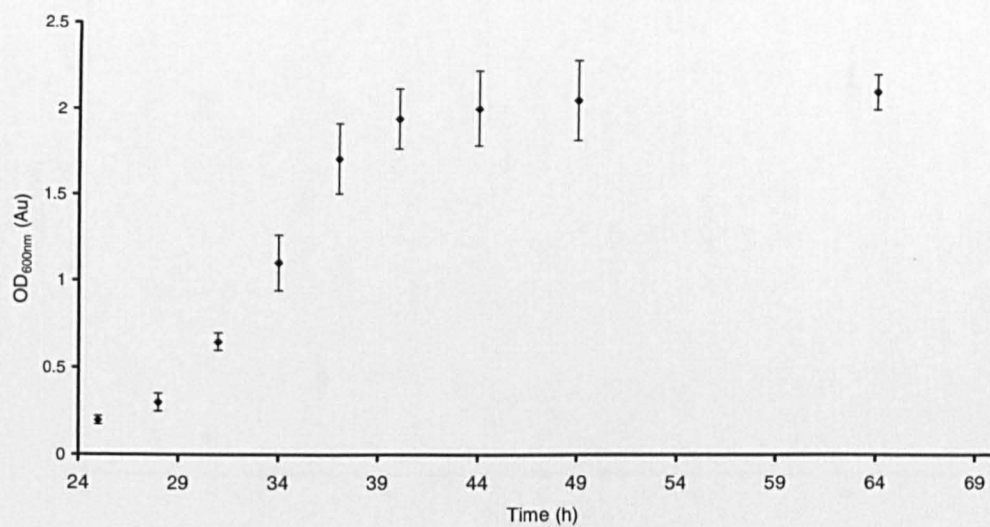


Fig 5.1a, *H. volcanii* growth in HvYPC. Optical density is plotted against time since inoculation. Error shown is standard error from 4 replicates.

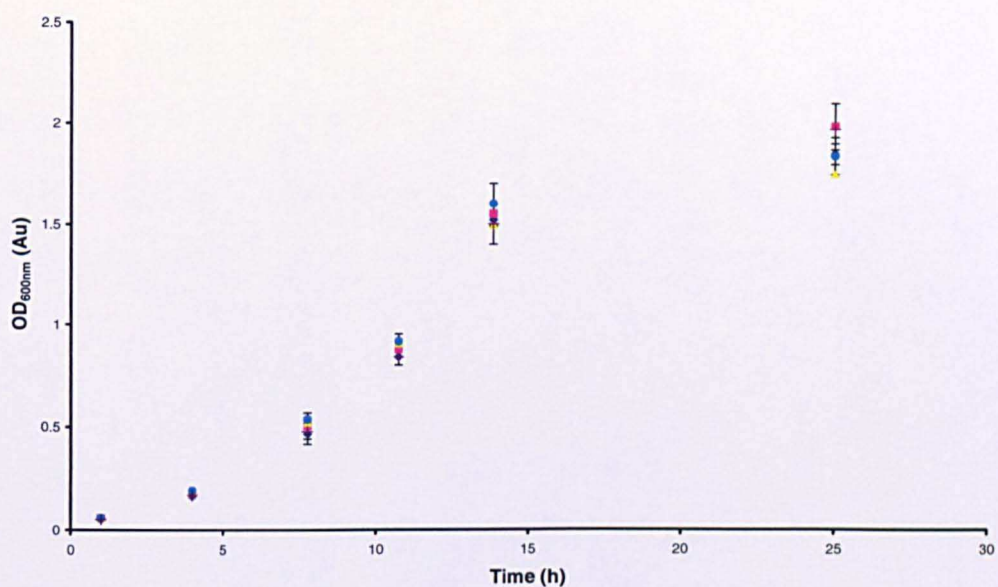


Fig 5.1b, *H. volcanii* growth in HvYPC. Optical density is plotted against time since the start of exponential phase. Circles show *H. volcanii* DS70, Diamond H26, Square H53, Cross  $\Delta TBP1$  and Triangle  $\Delta TBP3$

Table 5.1 Length of Lag Phase by Strain: This table shows the variation in lag phases in HvYPC of several *H. volcanii* strains determined on a strain by strain basis. Mean time followed by variance is shown for each strain, 4 replicates were used in each case.

Strain	Length of Lag Phase (h)
DS70	24 ± 4
H26	25 ± 4
H53	25 ± 3
<i>ΔTBP1</i>	36 ± 12
<i>ΔTBP3</i>	28 ± 5

### 5.3 Lycopene Production During Growth

That lycopene is produced by *H. volcanii* is a previously established fact (Bidle *et al.* 2007), however, when it is produced, how and in what quantities were unknown. For this reason it was decided to measure lycopene production throughout the exponential phase of *H. volcanii* growth, the phase in which the organism is most metabolically active. It was hypothesised that lycopene production would be seen to increase at a rate either accordant with the doubling of the cells or higher than this, as it had previously been established that lycopene was a highly energy expensive molecule to produce (Fig 3.1) it seemed reasonable to suggest that it would be produced during periods of energy abundance. In order to test this, an experiment was devised in which a sample would be taken every hour after a specific optical density, chosen to ensure that the culture had entered exponential phase. The selected optical density was 0.4 OD<sub>600</sub>, this point in the growth was selected as it could be repeatedly shown to be during early exponential phase (Fig 5.1), whilst still having enough cells present to be able to attain a measurable extraction of lycopene. The absolute mass of lycopene in the culture throughout exponential phase remains relatively constant (Fig 5.3a.), as the number of cells is increasing this means that the level of lycopene per cell is



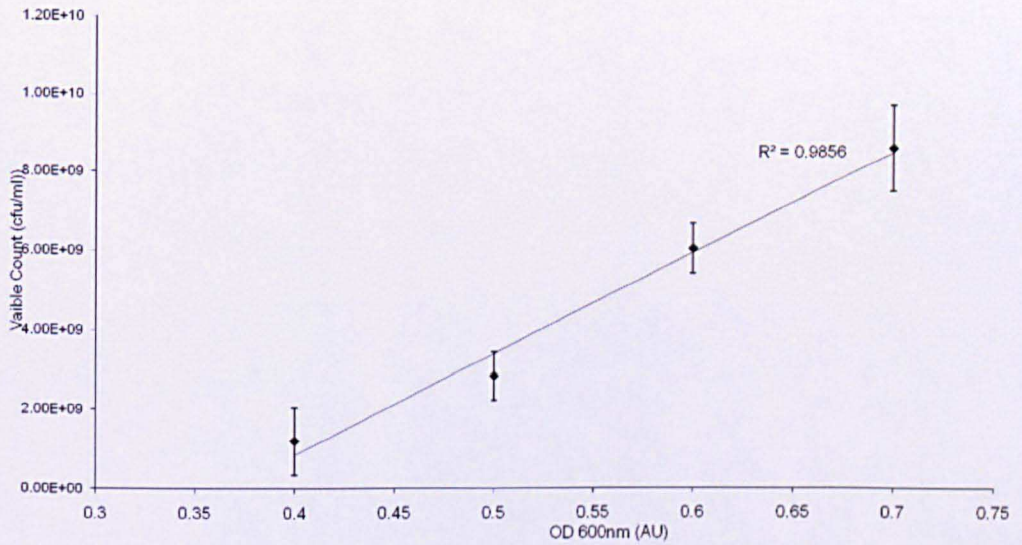


Fig 5.2a Optical Density in Au plotted against viable count in cfu/ml. Error shown is standard error from four replicates. Both scales are linear.

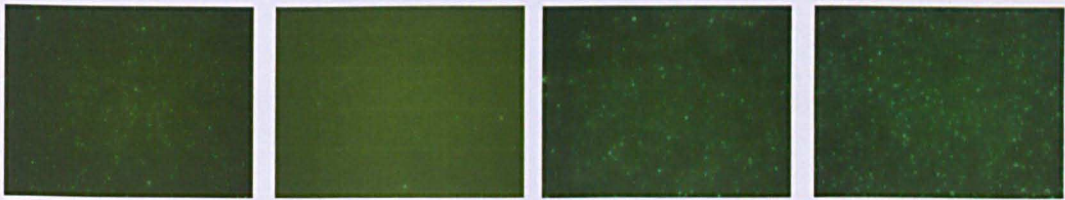


Fig 5.2b, Fluorescence images taken using Bac-Light showing Optical Densities 0.4, 0.5, 0.6 and 0.7 respectively, these images were used to calculate the number of cells in the culture

reducing (Fig 5.3b). The reduction of lycopene per cell throughout exponential growth suggested that production of the compound during this phase of growth is curtailed. Such a pattern of production is consistent with secondary metabolites, such as antibiotics (Takano *et al.* 1992). Many secondary metabolites are produced in stationary phase as a response to the stresses placed on the cell in these conditions. It is not unreasonable to think that lycopene may be used in such a manner as it is a strong antioxidant (Glauert *et al.* 2010) as well as a light absorbing compound capable of screening frequencies which may be damaging to proteins (Carolina 2004) (Fig 5.4). To discover if such production was indeed occurring lycopene levels in later stages of growth needed to be measured. To investigate the hypothesis that lycopene was produced in stationary phase at greater levels than in exponential phase an experiment was undertaken in which lycopene levels were measured both in exponential phase and stationary phase growth. This was performed by selection of a point in exponential phase as used previously and an optical density at which stationary phase growth was clearly occurring. It was decided by analysis of growth curves (Fig 5.1) that an optical density greater than 1.8 OD<sub>600</sub> would be a suitable indicator of stationary phase growth. In addition 0.4 OD<sub>600</sub> was once again used to sample exponential phase. The sample taken in exponential phase was checked for consistency against those taken in the exponential phase experiments at the same point to ensure that an

accurate characterisation was being made. The comparison between the levels of lycopene in exponential and stationary phase growth show a highly significant difference by t-test ( $p < 0.01$ ), with lycopene levels being higher in stationary phase growth (Fig 5.5.). The question that now presents itself is what the function of lycopene might be. The answer to this may lie in the nature of lycopene itself and therefore consideration of its properties as a light absorbing and chemically reductive compound are of great interest.

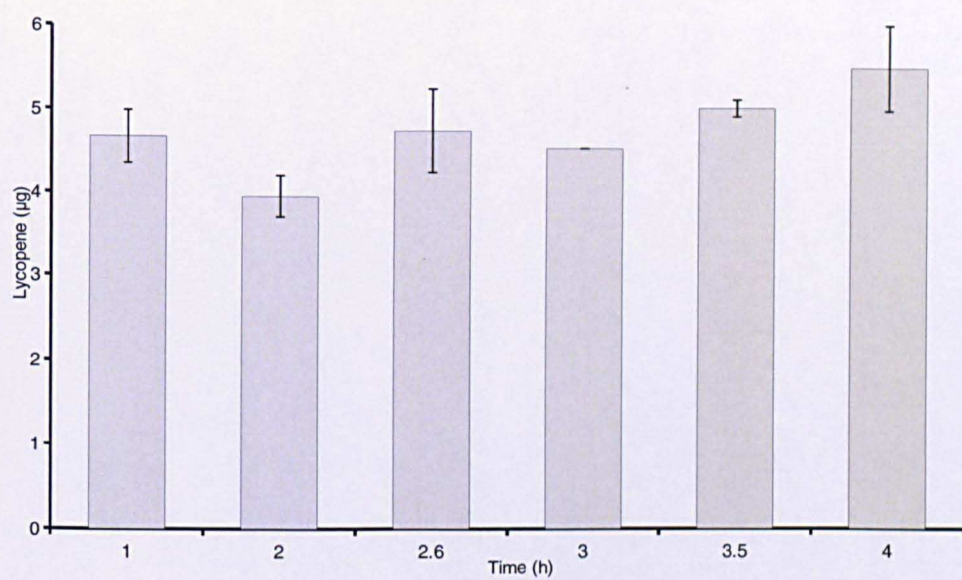


Fig 5.3a, Lycopene levels ( $\mu\text{g}$  of lycopene present in 200 ml culture) against time since start of exponential phase. Error bars show standard error from 4 replicates.



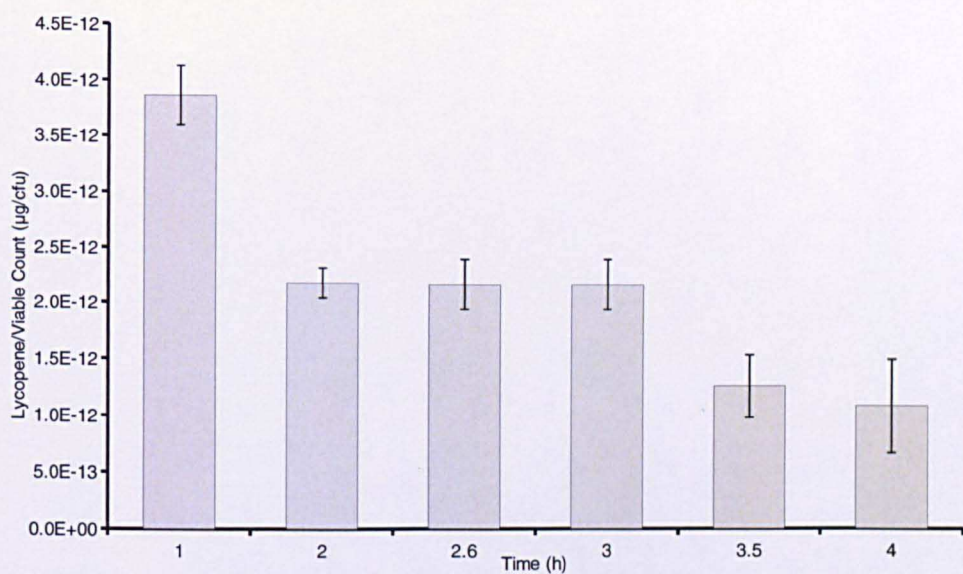


Fig 5.3b Lycopene per viable cell (μg/cfu) against time since start of exponential phase across exponential growth. Error bars show standard error from 4 replicates.

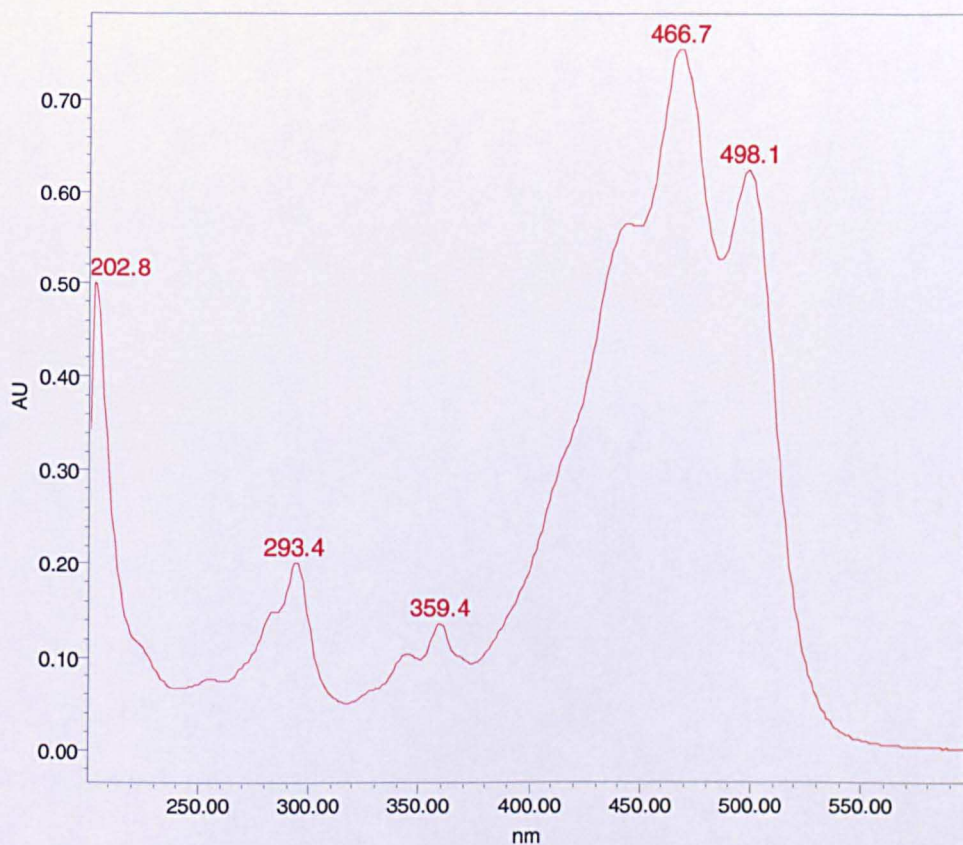


Fig 5.4 Spectrum of lycopene in 1:19 THF:MeOH showing absorbance against wavelength. This was obtained using the method for isocratic HPLC and a photodiode array detector as detailed in Chapter 2.

## 5.4 Lycopene Production Under Stress Conditions

It is not uncommon that a compound be produced by an organism only when necessary. For this reason it is possible to use an assessment of a response in production of a compound to assess its function. In this case it was decided that in order to determine if lycopene's light absorbing characteristics played a part in it being useful to the cell the production of lycopene in the dark and under illumination should be considered. To investigate this *H. volcanii* cultures were grown, 5 of these cultures exposed to light and 5 not, with sunlight during the day and electrical illumination by night to ensure the lighted half was stimulated and aluminium foil wrappings to keep the dark cultures from the light. Cultures were sampled across exponential phase and in stationary phase. The lycopene levels in each culture were then analysed and it was discovered that there was no significant difference between the light and the dark cultures (Table 5.2). It was therefore considered unlikely that lycopene functioned as a protective agent against light.

As lycopene is significantly reductive (Glauert *et al.* 2010) as well as light absorbing it was decided that an appropriate stress to expose the organism to

would be oxidative stress. In order to induce the stress an appropriate means of causing an increase in reactive oxygen species (ROS) in a controlled way is necessary. It was therefore decided that a controlled amount of hydrogen peroxide would be used in order to apply the necessary treatment. This, however, presented a new problem as HvYPC is a rich medium containing several reducing agents in uncontrolled amounts; it is therefore clearly unsuitable for use in an experiment in which the levels of ROS are critical. For this reason it was decided that a chemically defined medium was required.



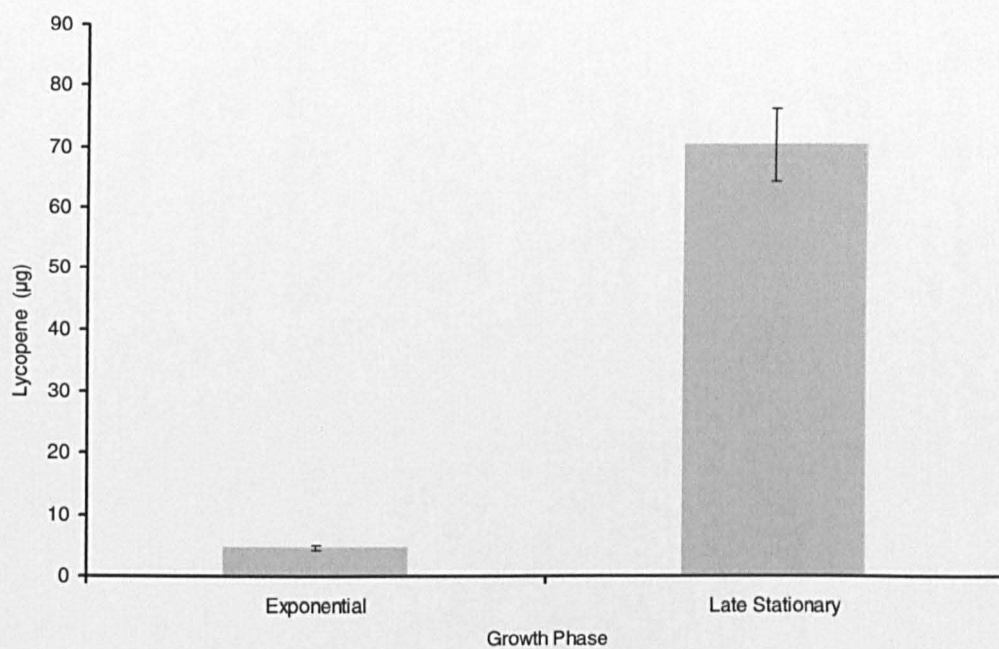


Fig 5.5, Average lycopene levels in mid exponential and stationary growth of *H. volcanii* shown as lycopene per 200 ml culture (μg). Error bars shown are standard error from 10 replicates.

## 5.5 Chemically Defined Media Development

Chemically defined media open up the possibility of performing certain experiments, such as those inducing chemical shock and removal of specific media components, to allow a more full analysis of the behaviour of the organism.

Pre-existing chemically defined media were noted to be none functional in liquid culture (although functional as solid medium). For this reason it was necessary to modify the basic media production in such a way as to allow the growth of the archaeon.

The base media used as starting point for development was Thorsten Allers HvMin. It was decided that, firstly, a single carbon source should be selected in order to minimise complicating factors in growth. As lactate is easily metabolised by many organisms it was decided that sodium lactate would be an appropriate carbon source. The media was then produced exactly as HvMin, with the mixed carbon source replaced with 20% sodium lactate. This still failed to show any growth and so it was necessary to consider what other changes might help to stimulate *H. volcanii*. Calcium chloride is an integral

Table 5.2 Lycopene levels in  $\mu\text{g}$  per 200ml culture in *H. volcanii* DS70 grown in light and dark conditions. Error shown is standard error from four replicates.

Time Since Innoculation (h)	Light ( $\mu\text{g}$ )	Dark ( $\mu\text{g}$ )
16	$3.9 \pm 0.3$	$4.2 \pm 0.4$
17	$4.1 \pm 0.2$	$4.1 \pm 0.5$
18	$3.9 \pm 0.4$	$4.4 \pm 0.4$
19	$4.3 \pm 0.3$	$4.2 \pm 0.2$
52	$63.2 \pm 8.1$	$61.5 \pm 10.2$

part of any *H. volcanii* medium, including that of HvYPC and for this reason an increased level of calcium chloride was considered to be a potential remedy to the lack of growth. When an extra 1 mmol of calcium chloride was added growth was observed. The media produced in this way was designated Hv Na-Lac. Successful production of a chemically defined medium allowed the planning of the desired oxidative shock experiments; it also, however, required that the new media be shown to work similarly to HvYPC to ensure that accurate conclusions were drawn.

For the purpose of comparing HvYPC with Hv Na-Lac a series of experiments were needed in which certain behaviours already characterised in HvYPC could be analysed in the new medium. Growth patterns in HvYPC and Hv Na-Lac were compared in order to establish that growth in the new media was similar. It was observed that, whilst in Hv Na-Lac a longer lag phase was typically observed the exponential and stationary phases proceeded on a time scale concordant with that of HvYPC (Fig 5.6a) for *H. volcanii* DS70. This suggested that the organism behaved similarly in both media. Other strains to be used in the subsequent experiments were then tested to assess their individual growth rates and establish if they were comparable to one another in the new medium. The strains used in this analysis were *H. volcanii* DS70, H53, H26 and  $\Delta$ TBP3. These were selected

as H26 is the most similar to the knock-out strain, whilst H53 is the mutants parent strain and DS70 the wild type isolate. It was shown that all strains showed similar growth in Hv Na-Lac (Fig 5.6b). This encouraged the use of Hv Na-Lac for subsequent experiments.

Lycopene production in Hv Na-Lac if different to that observed in Hv YPC would need to be reassessed for all previous experiments. The experiment assessing lycopene levels in exponential and stationary phase was repeated with Hv Na-Lac cultures, with the result that no significant difference was found between the levels in Hv Na-Lac and those in HvYPC, 4 replicates were performed and were found to be statistically similar ( $P>0.9$ ) by ANOVA at both growth phases. It was therefore concluded that for the purposes of this project both media caused behaviour of a highly similar nature during growth and therefore Hv Na-Lac could be used appropriately in the subsequent stress level experiments.

## **5.6 Carotenoid Production Under Oxidative Stress**

Lycopene's nature as an anti oxidant, capable of reducing ROS such that they no longer pose a danger, is a possible reason for its production by the cell. In order to

assess if this is indeed the case an experiment was designed whereby several cultures of *H. volcanii* would be set up in parallel with a group of controls and a separate group to be subjected to oxidative shock. From the literature available (Tian *et al.* 2007) it was determined that 100  $\mu$ M hydrogen peroxide would produce such a shock without proving lethal to most organisms. Accordingly the experimental group were treated with this concentration of hydrogen peroxide after a designated point in the growth curve. As with the earlier experiments in this section 0.4 OD<sub>600</sub> was chosen as this point, both for consistency and as changes due to shock were likely to be more readily detectable when levels of lycopene were low, as they are in exponential phase (Fig 5.5.). After treatment samples were taken every hour for four hours both from the control and experimental groups. The control showed no difference when compared to the previously observed lycopene levels for exponential phase growth. The experimental group, showed a slight, but clearly significant ( $P < 0.01$ ) difference. Across all time points the lycopene levels in the experimental group could be clearly seen to be lower than those of the control (Fig 5.7) This, it seems likely, is caused by a depletion of lycopene as it reacts to the hydrogen peroxide and, whilst indicative of an interaction between the stress and the compound, this does not constitute a reaction on the part of the cell.

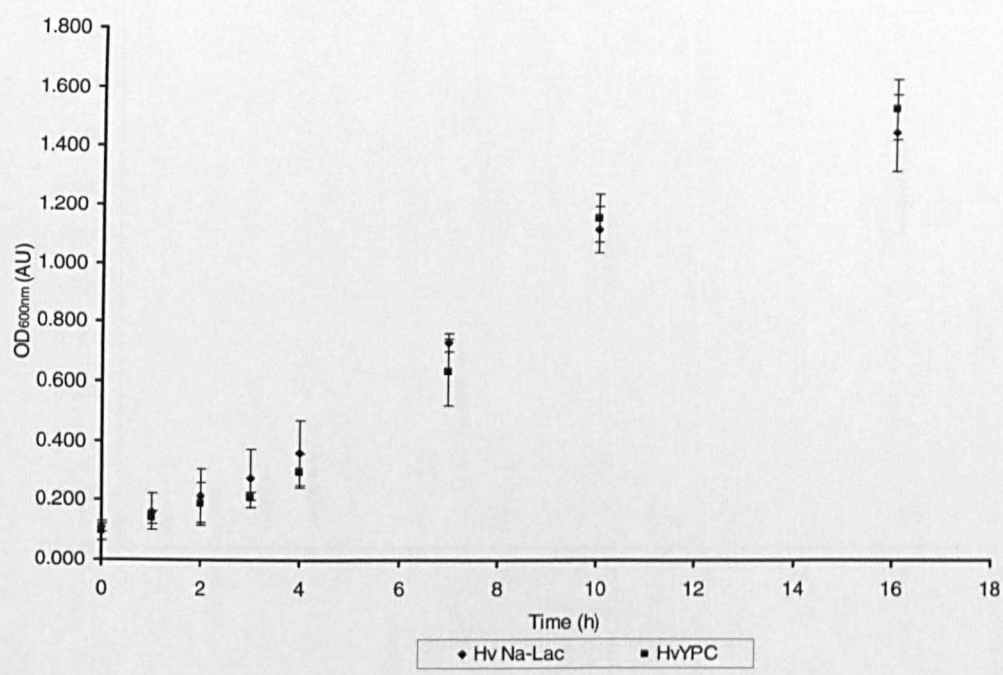


Fig 5.6a, Growth curves post lag phase of *H. volcanii* DS70 in both HvYPC and Hv Na-Lac

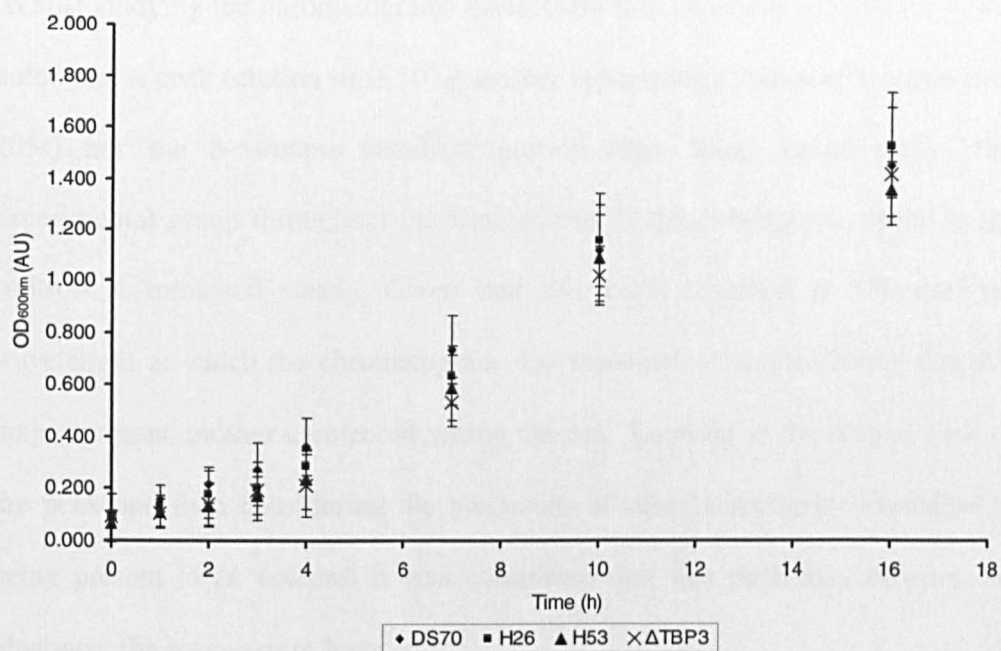


Fig 5.6b, Growth curves post lag phase of all strains in Hv Na-Lac. Optical density is plotted against time after the start of exponential phase. Errors shown are standard error from four replicates.



Whilst studying the chromatograms taken from this experiment, however it was noted that a peak (elution time 301s) neither representing lycopene (elution time 205s) nor the  $\beta$ -carotene standard (elution time 260s) varied across this experimental group throughout the time course of the experiment, whilst in the controls it remained steady. Given that this peak absorbed at 450 nm, the wavelength at which the chromatogram was recorded, it seemed likely that this may represent another carotenoid within the cell. Looking at the elution time of the peak and then considering the properties of other carotenoids identified as being present in *H. volcanii* it was considered that this peak may prove to be phytoene, the precursor to lycopene.

To assess this a gradient HPLC method with an in line photo diode array (PDA) detector was used (Section 2.12). Once the peak suspected to be phytoene was eluted the spectrum was obtained (Fig 5.8). This spectrum was compared to published values of carotenoids in several similar solvent systems (Britton *et al.* 1998) as the solvent system used (1:19 THF:MeOH) was not one in which the compound had been characterised. This analysis indicated that the eluant in question was indeed phytoene, which allowed the analysis of the data collected previously to be performed in light of this. The variation in phytoene levels under

oxidative stress (put in data showing stress response) was considered an important factor to analyse as variations in the precursor could compensate for variations in lycopene itself. This could be so if, for example, the cell had a pool of phytoene from which new lycopene could be quickly synthesised, thus compensating for the drop in the levels of lycopene. This would, however create noticeable effects in phytoene levels, such that they would be seen to decrease and, once the cell responded to the drop, then be seen to increase once more. In order to ensure that any increase or decrease in phytoene levels was genuine rather than an artefact of extraction the peak height of the phytoene peak was converted into a ratio between itself and the  $\beta$ -carotene standard, much as the peak height of the lycopene peak was prior to estimation of its true amount. Direct calculation of lycopene concentration using literature molar extinction coefficients were not suitable for this analysis as losses during extraction could cause variations in the results. Phytoene, unlike lycopene, is not readily available to use as a standard and therefore no calibration curve could be built to estimate the true amount of phytoene. Instead the absorbance ratio was used as an approximation of this and an indicator of phytoene levels. When compared to the ratios obtained from the control samples, which remained constant throughout exponential growth in a similar manner to lycopene, the phytoene levels in the experimental sample were seen to first decrease sharply before recovering to levels greater than their original

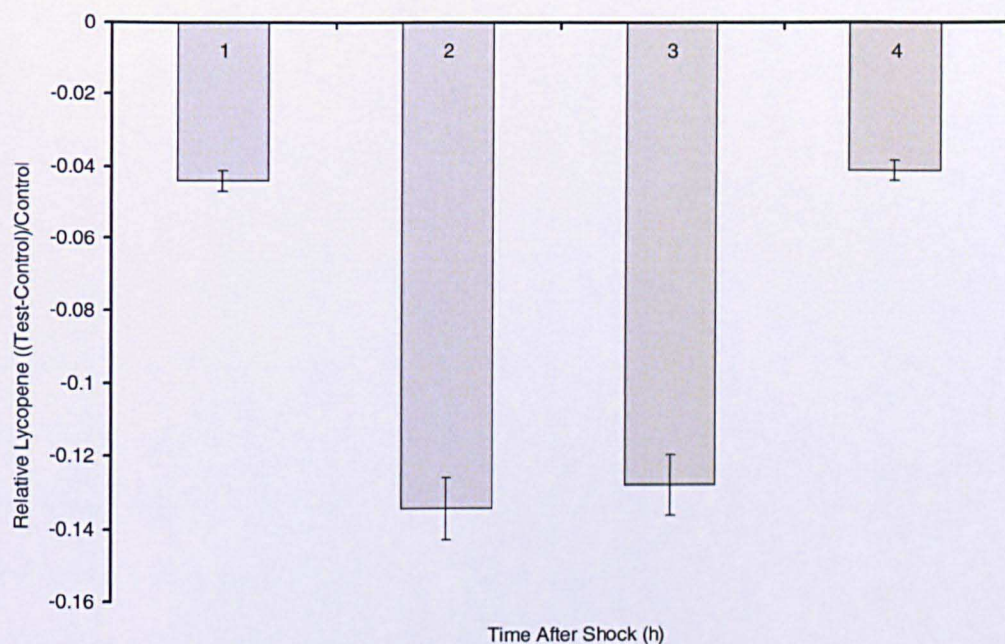


Fig 5.7, Change in lycopene levels after 100  $\mu$ M hydrogen peroxide treatment, showing relative lycopene against time since treatment. Error shown is standard error from 4 replicates.

base state (Fig 5.9a). This is likely due to lycopene levels dropping when attacked by hydrogen peroxide and then being replenished rapidly, drawing on the phytoene pool for a quick response.

It was hypothesised that should the cell be subjected to a strong enough oxidative shock it would be reasonable to think that lycopene destruction would increase to such an extent that the pool of phytoene acting as a buffer would be exhausted and lycopene then be depleted as a result. In order to investigate this, the prior experiment was repeated with an increased hydrogen peroxide concentration, 1 mM. This increased concentration caused no significant change in the behaviour of phytoene (Fig 5.9b), however lycopene was seen to decrease and then increase significantly ( $P < 0.01$  in both cases by ANOVA) in a similar manner to that observed for phytoene under both shock conditions (Fig 5.10).

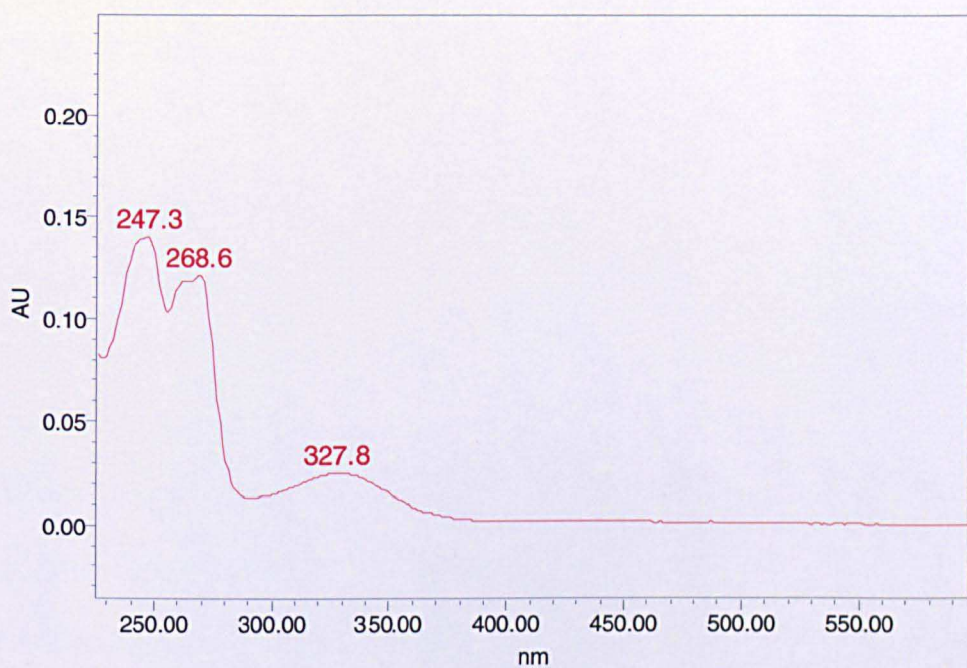


Fig 5.8, Spectrum of phytoene in 1:19 THF:MeOH showing absorbance (AU) against wavelength (nm).

## **5.7 Post Lycopene Carotenoid Products of *H. volcanii***

Whilst analysing the chromatograms prepared during interrogation of the oxidative shock experiments several peaks were noted prior to the elution time of lycopene (320s). These peaks at elution times 183s and 249s were initially thought to be degradation products. When the spectra of these peaks were elucidated (Fig 5.11 a and b) the spectra showed a clear three peak structure, characteristic of carotenoids, but different to that of lycopene. As these peaks appeared to progress towards the solvent front in a reasonably regular sequence it was decided to check the solvent front for any similar spectra showing three peaks and, when this was done, one was found. To isolate this from the background noise present in the solvent front due to its containing cell components that did not adhere to the column it was decided to alter the gradient of the solvent system, starting at 1:9 Water:MeOH and proceeding to 1:19 THF:MeOH over 30 minutes, in order to increase peak separation. This done a third and fourth carotenoid peak ahead of lycopene could be distinguished and a spectrum obtained.



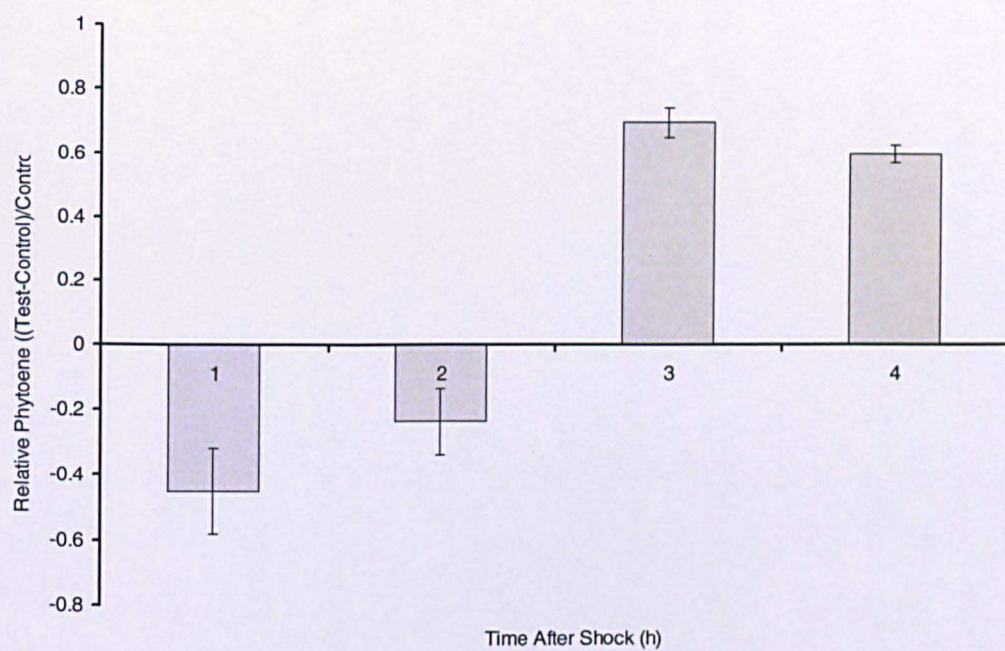


Fig 5.9a, Phytoene change relative to control against time after shock. Errors shown are standard error from 4 replicates.

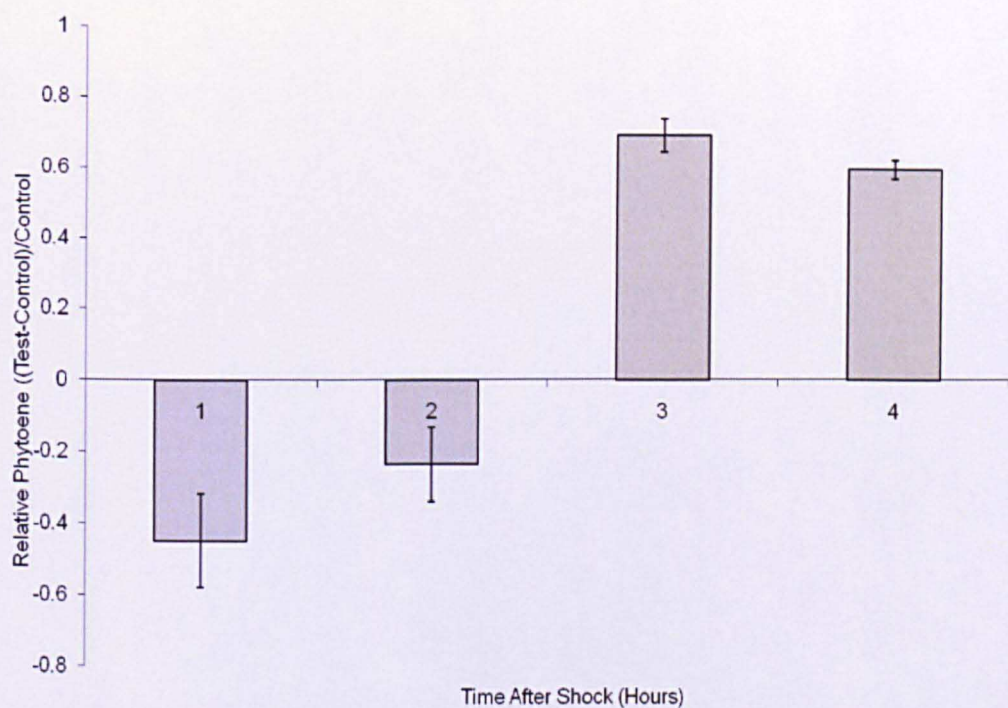


Fig 5.9b, Phytoene change relative to control against time after shock. Errors shown are standard error from 4 replicates.



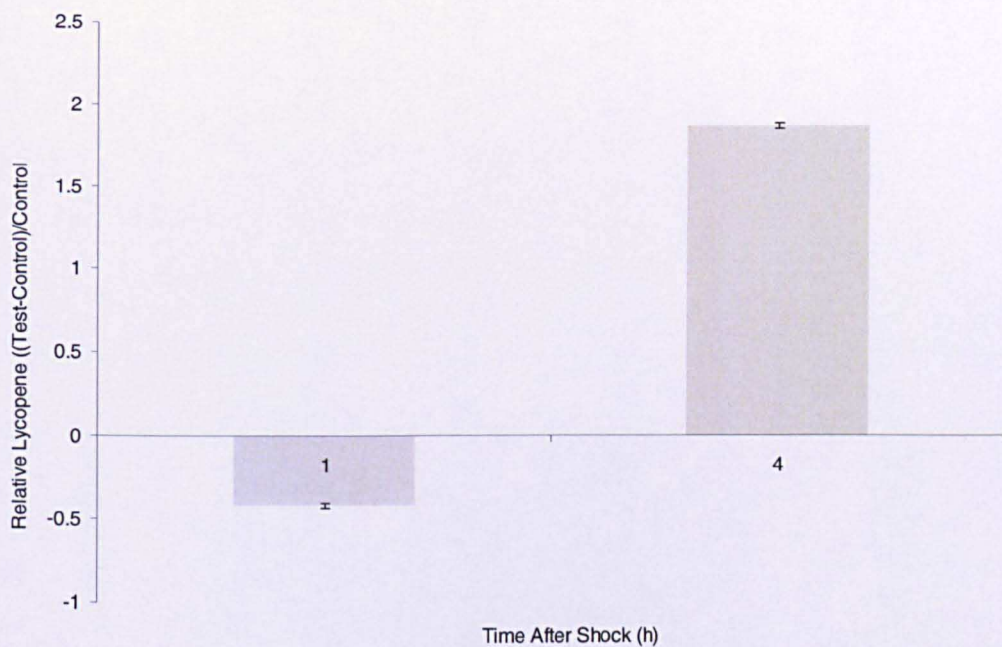


Fig 5.10, Lycopene changes relative to untreated control against time after treatment. Errors shown are standard error from 4 replicates.

It was decided that a comparison of the shift caused to peak absorbances of both lycopene and phytoene in the running solvent against pure methanol would be assessed. It was discovered that peak absorbance varied by around 5-20 nm above the spectrum in methanol in all cases, with greater differences at longer wavelengths. This information allowed for an assessment of carotenoid profiles in methanol to be used to show which of this family of compounds the extra peaks were.

Upon an interrogation of the literature spectra for carotenoids (Britton *et al.* 1998) it was discovered that the most likely candidate for peak four (Fig 11 c.) was likely to be bacterioruberin. As bacterioruberin is typically produced in a four step process by which similar hydrocarbons are added sequentially to each end of a lycopene molecule and oxygens then added (Schweitzer *et al.* 1966) it was suggested that the other unidentifiable compounds may well be these interim steps. The regular dispersal of the peaks on the chromatogram along with the increasing hydrophilicity, produced by added oxygen groups with each addition, would also fit this conclusion.

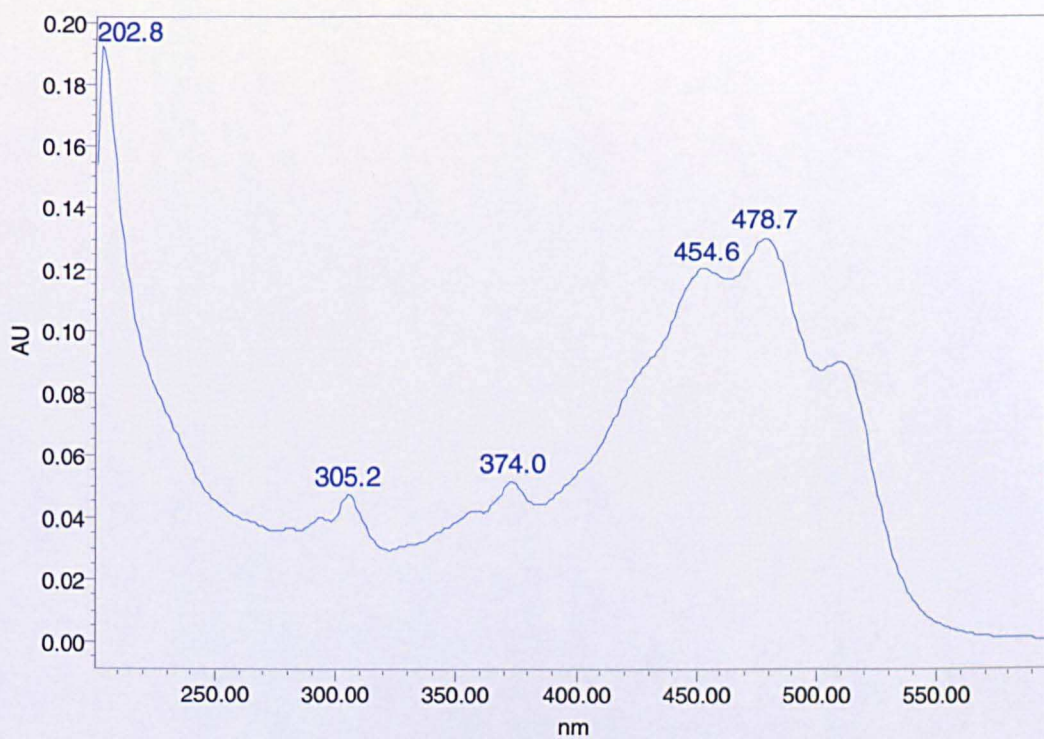


Fig 5.11a, from unknown peak 1, elution time 249s showing absorbance (AU) against wavelength (nm)



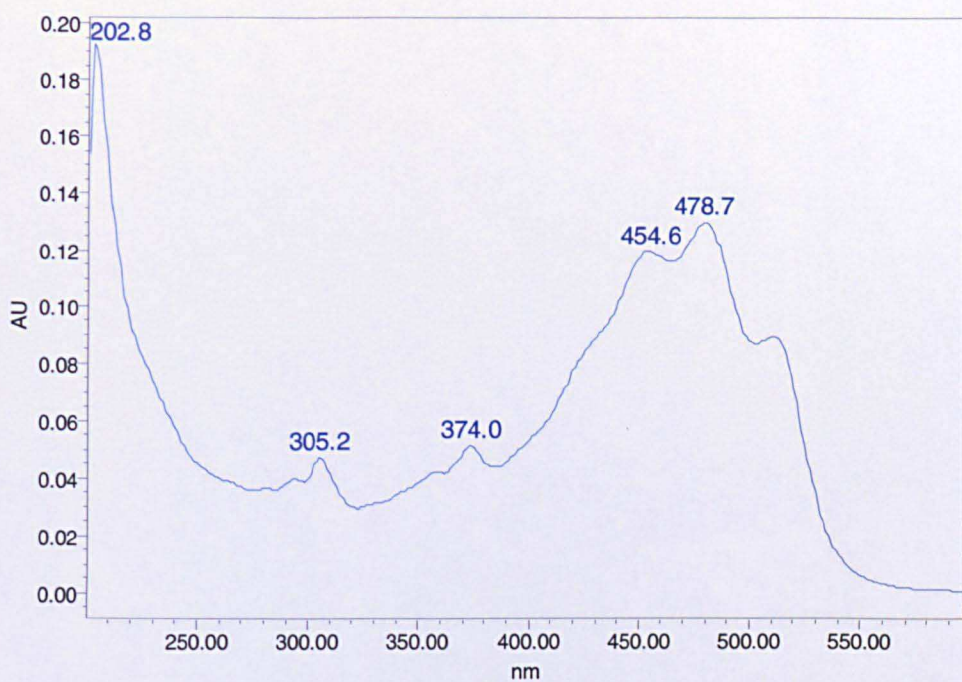


Fig 5.11b, Spectrum from unknown peak 2, elution time 180s showing absorbance (AU) against wavelength (nm)

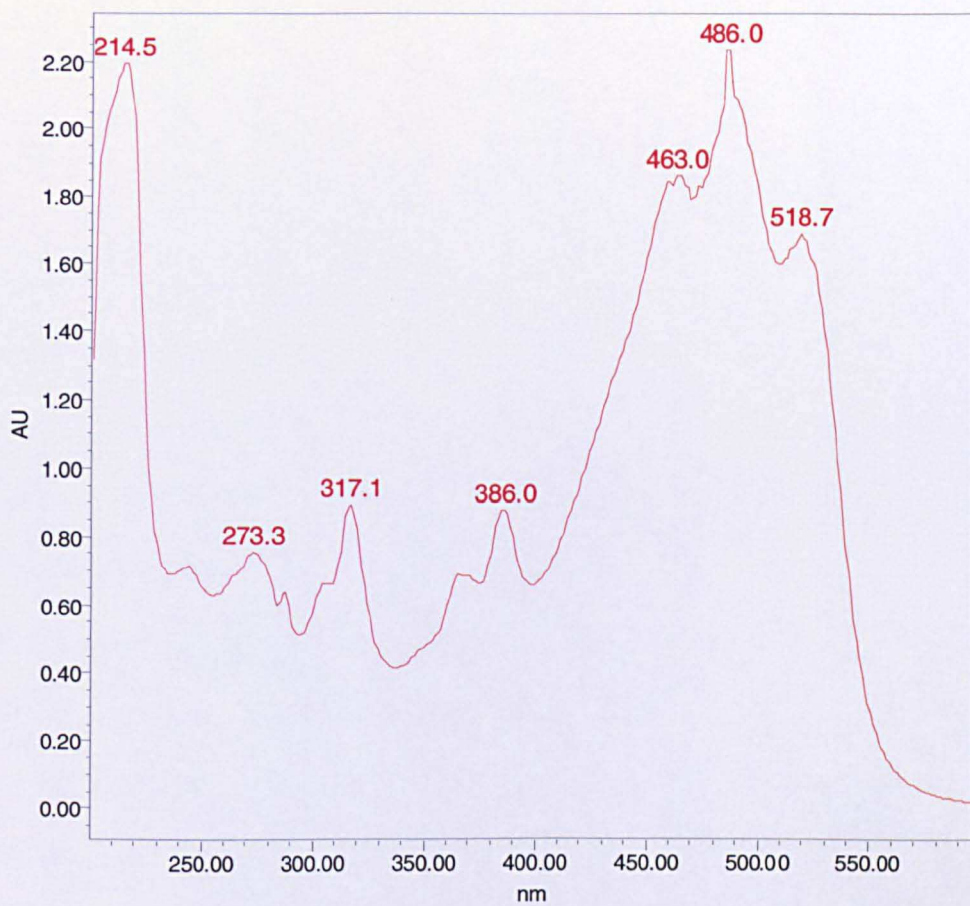


Fig 5.11c, Spectrum of Bacterioruberin showing absorbance (AU) against wavelength (nm)

## 5.8 Discussion

Once it was established that *H. volcanii* strains performed in our hands in a comparable way to the manner recorded in the literature it allowed for studies on lycopene production to commence. Whilst initial experiments were conducted in the commonly used Hv YPC later experiments requiring greater control over the components of the medium were then conducted in Hv Na-Lac a chemically defined medium developed for this project. Although a longer lag phase was typically observed in Hv Na-Lac similar growth post lag phase and similar lycopene content at all conditions tested were enough to assure us that no meaningful differences came from the altered culture conditions. It was discovered that lycopene levels on a per cell base decreased during exponential phase, counter to the initial hypothesis. When lycopene levels were sampled across a broader range of the growth curve it was then discovered that they were highest in the stationary phase of growth. In the stationary phase cell resources are stretched, as such fewer non essential genes are active (Brock 2008; Gregor & Pfeiffer 2005) and lycopene production is an extremely energy intensive process, using 24 ATP and 12 NADH per molecule (Fig 3.1.), suggesting that it is important for cell viability in stationary phase.

As the role of lycopene within the cell has not been previously established, it cannot be claimed with any certainty that it is a necessary compound for cell viability, however.

It could, given the function of lycopene as a precursor to other carotenoids be simply an artefact, a pathway remaining in the cell subsequent to its later, functional part being lost.

The high energy requirement coupled with its production during the most energy deficient phase of the cells life, however, indicates that it is unlikely to be an artefact of a former cellular process, as this would then favour cells which became lycopene deficient and would likely have rapidly encouraged the removal of the pathway from the organism due to the pressures of competition. If it is, as seems likely, not an artefact it, or its product, bacterioruberin must logically be a functional compound within the cell.

As lycopene has several important properties which may be utilised by the cell the determination of a likely function was not a trivial task. The most likely roles that lycopene might play in the furtherance of the cells function were a protective

agent, either against light or oxidative shock, given its high proportion in the membrane of the cell and extreme hydrophobicity. These two possibilities lead to the logical assumption that lycopene levels will respond to light levels or oxidative stress. Over the course of this chapter I have shown that lycopene in none mutant *H. volcanii* does indeed respond to oxidative stress, although not to light exposure. Whilst lycopene levels simply drop in response to low level oxidative damage, a reasonable occurrence regardless of function as ROS degrade lycopene, it is the cells response to high levels of oxidative attack that provide the most interesting results. The response of the cell to oxidative attack is to increase lycopene and phytoene production rapidly, scaling production to such a degree that the level of lycopene quickly exceeds base levels. This response suggests that the cell does indeed require lycopene in these circumstances as a protective agent.

It is not, however, conclusive that the cell responds to the stress and not simply to the depletion of lycopene. The lack of response to other forms of stress, however, does indicate that lycopene is not up regulated by the cells generalized stress response. This is interesting in light of its increase in stationary phase growth, as many changes that occur at this time are due to the generalized stress response of the cell to the conditions (Kultz 2003). It must therefore be considered that lycopene production is being up regulated at this phase of growth independently



of the global response and in response to a specific trigger, likely either nutrient depletion or the increased levels of ROS produced by the organism during this period. Until recently it was commonly thought that the production of ROS in cells would slow during stationary phase, due to a reduction in metabolism. It now seems, however, that this is not the case (Nystrom 2004). In stationary phase growth it appears common for organisms to undergo an increased oxidative stress, with oxygen removal often being able to alleviate loss of viability in stationary phase (Nystrom 2004). The requirement for protective systems against oxidative damage in other organisms for stationary phase survival suggests that whatever means *H. volcanii* uses to protect itself from reactive oxygen species would also need to be up regulated during this phase of growth, particularly during early stationary phase, when it is suggested that errors in protein production and misfolding may lead to an oxidative burst (Nystrom 2004).

A further possibility is that bacterioruberin provides the protective effects noticed and that the change in levels of lycopene may be in response to bacterioruberin production, much as the changes in phytoene levels are suggested to be in response to lycopene production.

In conclusion the results of this chapter indicate that lycopene protects the cell

from oxidative attack and suggest that the cell up regulates the production of this compound in these circumstances, they do not, however prove this.

## **Chapter 6**

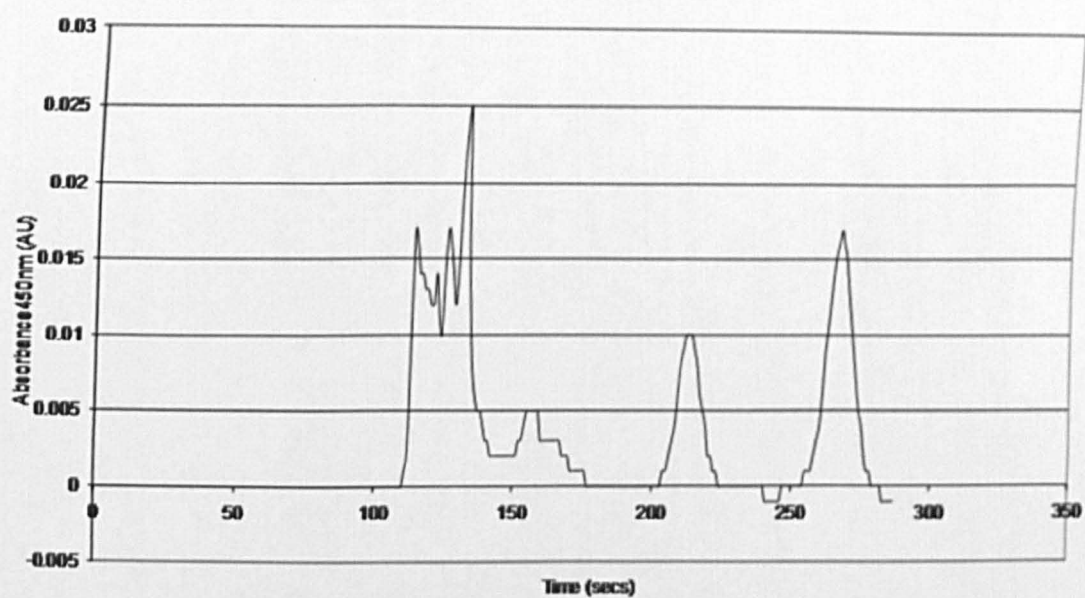
# **Construction of Gene Knockouts and Cloning for Heterologous Protein Expression**

## **Chapter 6 Construction of Gene Knockouts and Cloning for Heterologous Protein Expression**

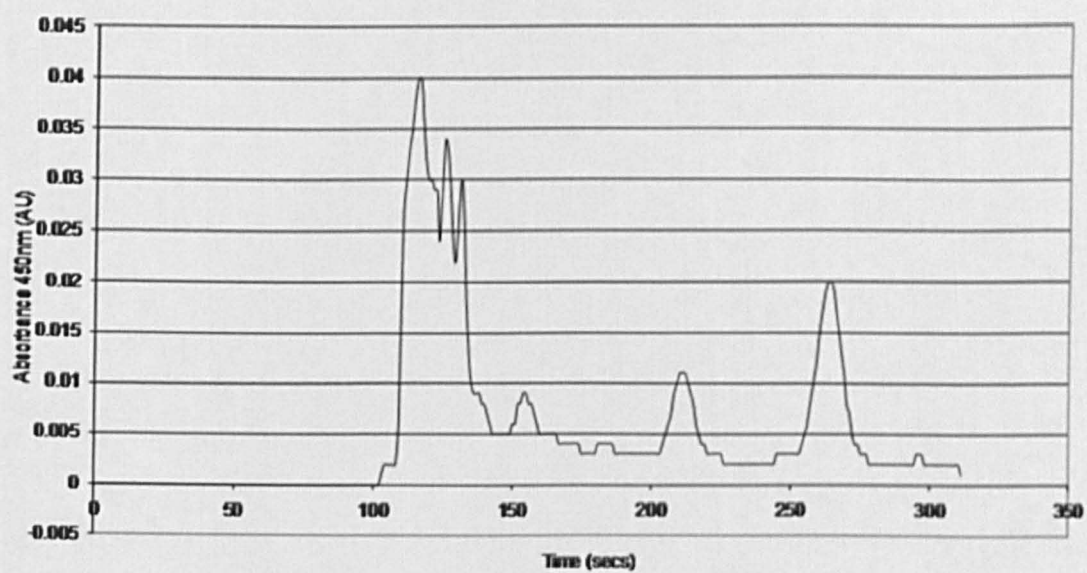
### **6.1 Introduction**

Whilst analysis of the behaviour of wild type organisms is important when studying the function of a particular compound produced by a cell it can only tell us so much. To be able to more fully understand the role of a particular compound it is useful to compare wild type behaviour to that of organisms deficient in the compound of interest, in this case lycopene. Working with an organism deficient in the compound of interest has several advantages, primarily it allows comparison of the wild type and the deficient strain in order to gain insights into the function of the compound. Experiments of this type allow the comparison of environmental factors that may affect viability in each strain such as oxidative stress, heat and light levels. In addition it allows the comparisons of each strain under normal growth conditions, which indicates at what phase of growth the compound of interest is of particular importance.

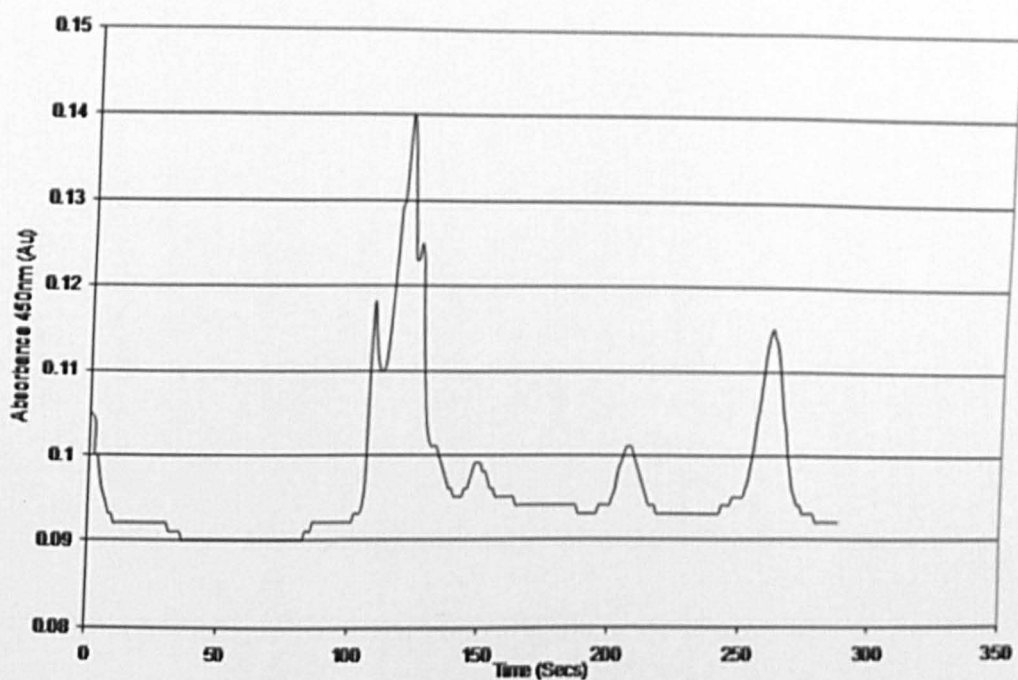
In order to achieve this it is necessary to produce organisms deficient in the compound of interest. There are a number of techniques for achieving this end. A common technique used, particularly for those traits which are phenotypically easily identifiable, such as colour, is to bombard the organism



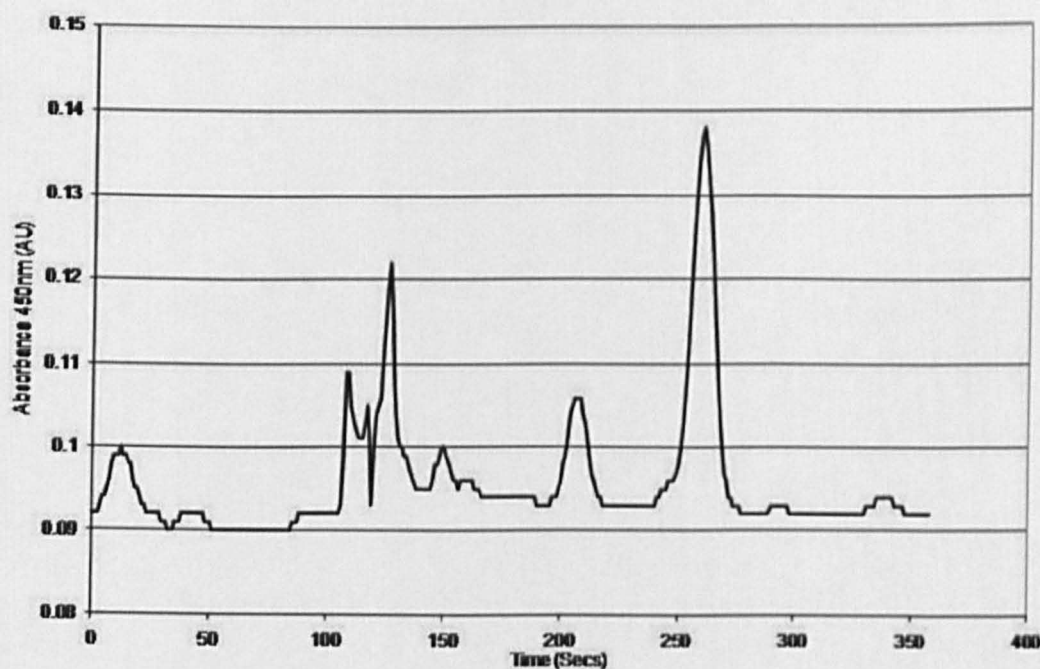
A



B



C



D

Fig 6.1. Chromatograms showing lycopene levels in wildtype DS70 (a) and white strains (b,c,d). Y axis absorbance at 450 nm (AU), X axis Time (s). Peaks at 200-225 seconds are lycopene, those at 245-275 are  $\beta$ -carotene

with ultraviolet radiation to cause mutations. UV radiation causes several different forms of mutation, by inducing strand breaks and thymidine dimers it can cause deletions, frame shifts and substitution mutations. All of these mutations can produce organisms deficient in a particular compound, thus simple exposure to UV followed by thorough screening of the viable progeny of the irradiated strains is all that is necessary. Particularly in the case of carotenoid production the selection of candidate strains for screening is a simple task as all that is required is to select those strains which present a white phenotype and assess their carotenoid content. When this technique is used for phenotypes that are more difficult to detect it can become highly labour intensive due to the level of screening required, but for the purposes of this project it is a reasonable method.

The downsides of random mutation induced by UV bombardment are the possible induction of mutations in genes other than those of interest. If this occurs it can lead to results being due to unintended mutations rather than those that are desired. Unintended mutations are difficult to screen for, as it would require screening the entire genome, it can therefore be appropriate to use a directed method for producing these mutations.

Targeted gene knockouts are such a method which can selectively remove specific genes from the organism without affecting the rest of the genome negatively. In general these methods involve homologous recombination, which removes the target gene from the genome and either replaces it with an

alternate gene, usually a marker, or a blank stretch of DNA. Typically systems which substitute a marker gene are superior as they allow selectivity for knockouts rather than simply for organisms that have taken up the knockout vector. Such a system is available for *H. volcanii*, produced by Dr. Thorsten Allers, University of Nottingham, which is a candidate for use in this study.

## 6.2 UV-Induced Gene Knockouts

There are several possible sources for mutants deficient in the production of lycopene. When an organism is bombarded with ultra violet (UV) radiation this often results in mutations in several of its genes. A collection of *H. volcanii* strains produced in this way was kindly made available by Dr. Thorsten Allers University of Nottingham. From these strains several that presented a white phenotype when examined on plates were selected. It was thought that these may be deficient in carotenoid production and could be used for comparative studies with wildtype strain.

To ensure the strains (H506, H514 and H529) were in fact lycopene deficient they were grown in HvYPC before carotenoids were extracted (Section 2.9-2.12). HPLC analysis of the extract showed that these strains, despite having a white appearance on a plate, did contain lycopene (Fig 6.1). Additionally they showed a pale red phenotype when grown in liquid culture. It was noted, however that the mutant strains had reduced lycopene levels compared to the wildtype strain DS70 (Table 6.1). From this it was concluded that, whilst



Table 6.1, Extraction Yields of UV Induced *H. volcanii* Mutant Strains

Compared to Wildtype. Yields given are in  $\mu\text{g}$  per 250 ml stationary phase culture OD600nm 2.0, 52 hours post inoculation. Standard errors calculated from three replicates.

Strain	Yield ( $\mu\text{g}$ )
H506	$66.2 \pm 1.6$
H514	$63.0 \pm 3.4$
H529	$58.4 \pm 2.0$
DS70	$70.2 \pm 1.1$

lycopene production may have been reduced, it had not been eliminated. As the reason for this down regulation was not evident, it could have arisen from a mutation in a large number of different pathways; these strains were not used in further analysis of lycopene production in *H. volcanii*.

### 6.3 Targeted Gene Knockouts

As opposed to UV bombardment, which produces random mutations, targeted gene knockouts seek to selectively remove specific genes from a genome. To do this they first require amplification of the regions of DNA either side of the gene to be knocked out. The gene was first located by searching an electronic copy of the genome and an area of around 1500 bp was copied from each side of the gene into a text file. Primer sites were then located to allow the insertion of restriction digest sites compatible with the plasmid vector to be used and gain an adequate yield of PCR product. The genes selected for this were *crtI*, *crtb2*, *pmdc* and *mvk*. They were selected due to their potential significance in the biosynthetic pathway (Fig 3.1). The restriction enzyme sites used in these primer pairs were KpnI and BamHI for the 5' flank, with BamHI being the downstream primer and SpeI and BamHI for the 3' flank with BamHI being in the upstream primer. Once PCR primers had been designed (Table 2.2) the next step was to acquire PCR product.

PCRs were run using all designed primer pairs with genomic DNA extracted from *H. volcanii*. Each set of primers was run on a temperature gradient, in

order to determine the best annealing temperature to use, with the brightest single band selected, *pmdc* 5' flank and *mvk* 3' flank showed no bands at any of the annealing temperatures used and therefore the primers for these were redesigned. The annealing temperatures for *crtb2* and *crtI* 3' and 5' flanks were selected (Table 2.2) and a DMSO gradient was used in an attempt to improve yield of PCR product without greatly increasing mispriming. The DMSO gradients were analysed by selecting the brightest band single band from an agarose gel (Section 2.18) on which PCRs performed with DMSO concentrations from 0% to 40% final PCR volume were run. This analysis showed high concentrations of DMSO gave the best results (Table 6.2) for these PCRs. This is not uncommon for PCR products produced from *H. volcanii*, due to its high genomic GC content. The identification of optimal running conditions for the PCRs allowed for larger volumes to be used in an attempt to secure a higher yield for cloning.

The knockout procedure requires a plasmid vector containing the flanks cloned from beside the gene of interest, a *TrpA* gene between these flanks and a *PyrE2* gene to be produced. The plasmid construct pTA131 (Section 2.3) contained *PyrE2* and a multiple cloning site into which the PCR products and the *TrpA* gene (with promoter) can be inserted. The PCR products were digested in double digests (Section 2.15) using FastDigest enzymes and buffer (Fermentas, York UK). The vector was digested in identical conditions and the digestions appeared to go to completion (Fig 6.2). The digested vectors were treated with Antarctic phosphatase (Promega, Southampton UK) and

purified before ligation was performed with the appropriate flanks (Section 2.16) Both the 5' and 3' flanks for both *crtI* and *crtb2* were treated in this way, as it was decided that if both flanks were successfully cloned it would be simpler to cut a flank from a plasmid for ligation than to attempt to cut To determine if the plasmid vector recovered in the attempt to make the knockout vector were in fact religations with mismatched digestion sites a digestion analysis was carried out. This required digesting the vector with the enzymes used in the initial digestion to check if the sites were still intact as well as digesting the vector with enzymes that would give a recognisable pattern confirming the identity of the plasmid. This analysis showed that the plasmid had indeed religated with mismatched sites as the identification was consistent with the plasmid, whilst digestion with KpnI and SpeI failed entirely. As religated vector was being recovered it was postulated that the problem may be in the digestion of the PCR products, as the successful digestion of these is impossible to tell by gel analysis, due to the fact that the removed DNA fragments make only a very minor difference in the size of the PCR product so a change in size is not readily detectable.

The PCR products were treated with a T4 polynucleotide kinase (Promega, Southampton UK) before they were ligated into longer pieces of DNA formed from multiple PCR products, more amenable to digestion by restriction endonucleases. In addition to making the sites more easily digested the longer DNA fragments were readily identifiable on an agarose gel and size changes were clear, showing successful digestion much as with a plasmid. The ligated

Table 6.2 Optimal DMSO Concentrations. DMSO concentration is represented by percentage of final PCR volume. Optimal concentrations were determined by selection of the brightest single band on an agarose gel on which PCRs performed with DMSO concentrations from 0% to 40% final PCR volume were run.

PCR	DMSO Concentration (%)
<i>crtI</i> 5' Flank	20
<i>crtI</i> 3' Flank	20
<i>crtb2</i> 5' Flank	15
<i>crtb2</i> 3' Flank	10

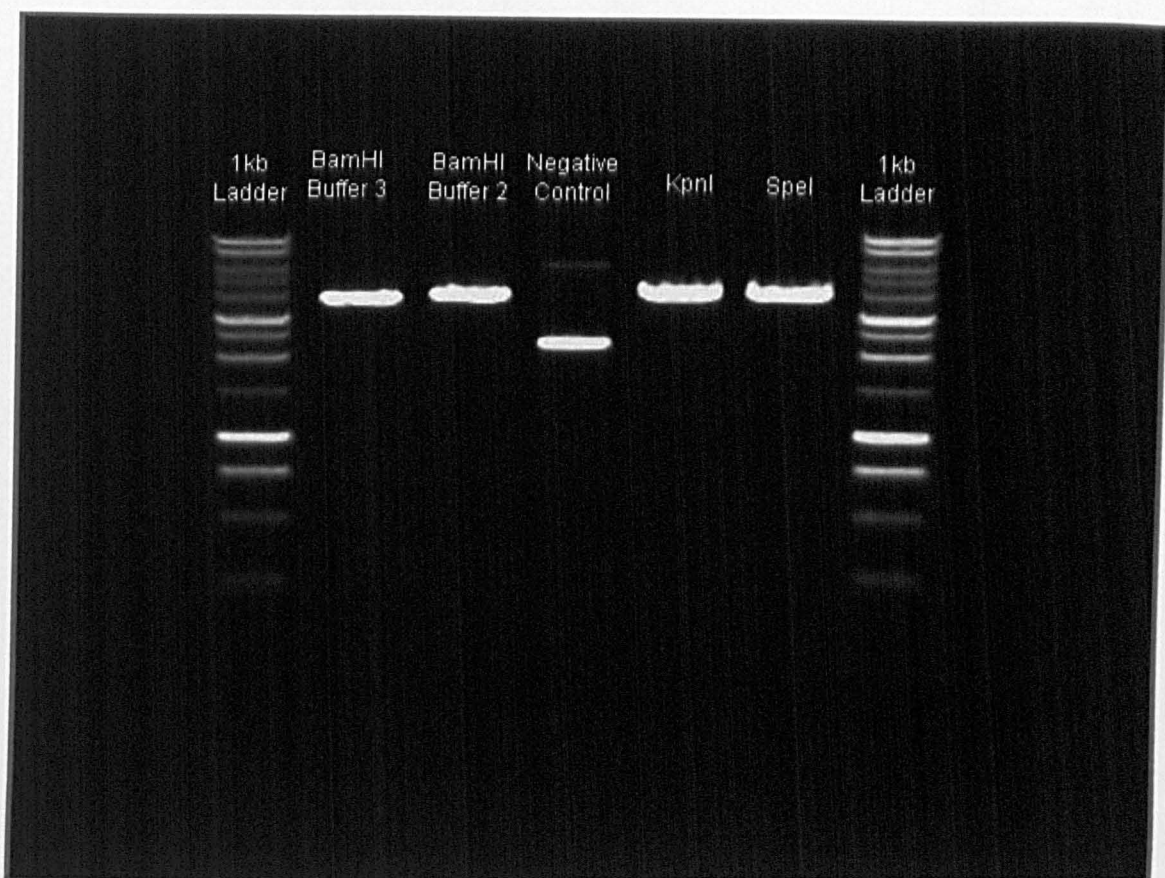


Fig 6.2.Digestion of pTA131 with BamHI, KpnI and SpeI, lanes 1 and 7 are promega 1 kb ladder, lanes 2 and 3 are pTA131 cut with BamHI in buffer 3 and 2 respectively, lane 4 is undigested pTA131, lane 5 is pTA131 cut with KpnI, lane 6 is pTA131 cut with SpeI

Table 6.3. Transformation efficiency of ligated plasmids. Transformation efficiency is given in colony forming units per ml.

Ligation	Transformation Efficiency (Cfu/ml)
pTA131 + <i>crtI</i> 5' Flank	70
pTA131 + <i>crtI</i> 3' Flank	25
pTA131 + <i>crtb2</i> 5' Flank	100
pTA131 + <i>crtb2</i> 3' Flank	30000



PCR products were digested with enzymes appropriate to the primers used in the PCR. The fragments of digested DNA were then gel purified (Section 2.18) for use in production of the plasmid vector for the knockout.

The plasmid containing the *crtI* 5' flank (pDR001) was digested such that the *crtI* 3' flank could be inserted. It was then treated with antarctic phosphatase for double the standard time before heat inactivation and purification. It was then added into a ligation reaction with the *crtI* 3' flank overnight on ice. Once the ligation reaction was completed the resulting plasmid was used to transform *E. coli* Top10 cells. Recovery was 15000 cfu/μgml and upon digestion analysis it was determined that both flanks appeared to have been successfully inserted. Digestion analysis alone, however does not check for exactly what DNA is present, only it's size. It was therefore possible that fragments of DNA of similar size to the flanks, or with mutations making them unfit for purpose had been inserted.

Screening the plasmids for incorrect insertions was performed by direct sequencing of the DNA insert using T7 primers, the sites for which are present in the vector. Sequencing results confirmed that the construct had been created successfully and this plasmid (pDR002) was taken forward to the introduction of the *TrpA* gene required for selection of knockouts.

The *TrpA* gene did not need to be cloned directly from *H. volcanii* genomic DNA as it was already available with appropriate restriction digest sites in

pTA298. pTA298 and pDR002 were digested with BamHI. pDR002 was then treated with Antarctic phosphatase, which was then heat inactivated, the resulting cut vector was purified. The *TrpA* gene and cut pDR002 were then ligated together and transformed. The resulting plasmid was checked by sequencing and digestion analysis, both of which confirmed it as the correct plasmid, designated pDR003. The successful production of the knock-out vector allowed the knock out procedure to be taken forward to *H. volcanii* transformation (Fig 6.3). *H. volcanii* strain H53 were utilised for mutant making as they contain deletions in both the *trpA* and *pyrE2* genes, allowing tryptophan and uracil auxotrophy and complementation to be used as selection markers. Transformation was carried out by first stripping the cells of their S-layer, causing them to become spheroplasts. These are delicate and require careful handling. The spheroplasts were then treated with EDTA to cause them to become susceptible to transformation. PDR003, containing a copy of the *pyrE2* gene in the backbone and *trpA* between the target gene flanks (Fig 6.3 H) was then introduced to the cells (Fig 6.4A) before the transformants were plated on media deficient in tryptophan, but containing uracil, as all cells would be uracil auxotrophic, whilst unmutated cells would be tryptophan auxotrophic. As *H. volcanii* undergoes homologous recombination frequently (Fig 6.4B) integrants would spontaneously occur. Colonies were selected from these plates and grown in rich liquid culture to cure the plasmid, ensuring that only copies resident on the genome would be present (Fig 6.4C). Subsequently they were spread on selective media containing 5-fluoro orotic acid 5-FOA and uracil, but no tryptophan, to select for double crossover events in which

the native gene had been lost (Fig 6.4Di), those cells still containing the pDR003 backbone, either as a genomic copy (Fig 6.4C) or as a free plasmid would be killed by production of 5-fluoruracil, whilst revertants to wild type (Fig 6.4Dii) would still be tryptophan auxotrophic and thus be unable to grow. They were incubated at 45°C for 5 days. At the end of this period no growth was present and the plates were stored at room temperature. After six months of storage at room temperature it was observed that a white lawn had grown on the plates previously thought to show no growth. This lawn of archaea was used as a target for PCR using primers designed to amplify the *crtI* region of the genome (Table 2.2). The primers showed a clear difference in length, which matches the predicted difference of 700 base pairs based on the difference between the length of the *crtI* ORF and the *trpA* containing insert, when compared to controls amplified from H53 genomic DNA (Fig 6.5). Additionally sequencing of these products showed a match to the predicted sequence for the knockout (Fig 6.6). The predicted sequence was generated by manipulating the plasmids and haloferax genome in silico, using the Clone Manager program. The pDR003 plasmid was generated using the same digestion and ligations as in the laboratory method. This was then manually inserted into the *H. volcanii* genome by deleting the regions homologous to those in the plasmid construct and the bases in between. The construct fragment consisting of the *crtI* flanks and *trpA* region was then copied into the gap. This generated two possible reference sequences, dependant on the orientation of the *trpA* region. When these were aligned to the sequence from

the PCR product the sequence depicted was found to match, the alternative sequence did not as the *trpA* cassette was in the opposite orientation.

The redesigned primers for the 5' and 3' flank of *pmdc* were found to produce suitable PCR yields for cloning. They were used to produce plasmid pDR006 in the same manner that the flanks of *crtI* were used to produce pDR003. Once successful production of pDR006 was confirmed by sequencing it was transformed into *H. volcanii*. No colonies were recovered on any of the subsequent plates, however leading to the conclusion that *pmdc* knockouts are likely lethal, indicating that the mevalonate pathway is likely the only isopentyl pyrophosphate (IPP) producing pathway in *H. volcanii*, which is in

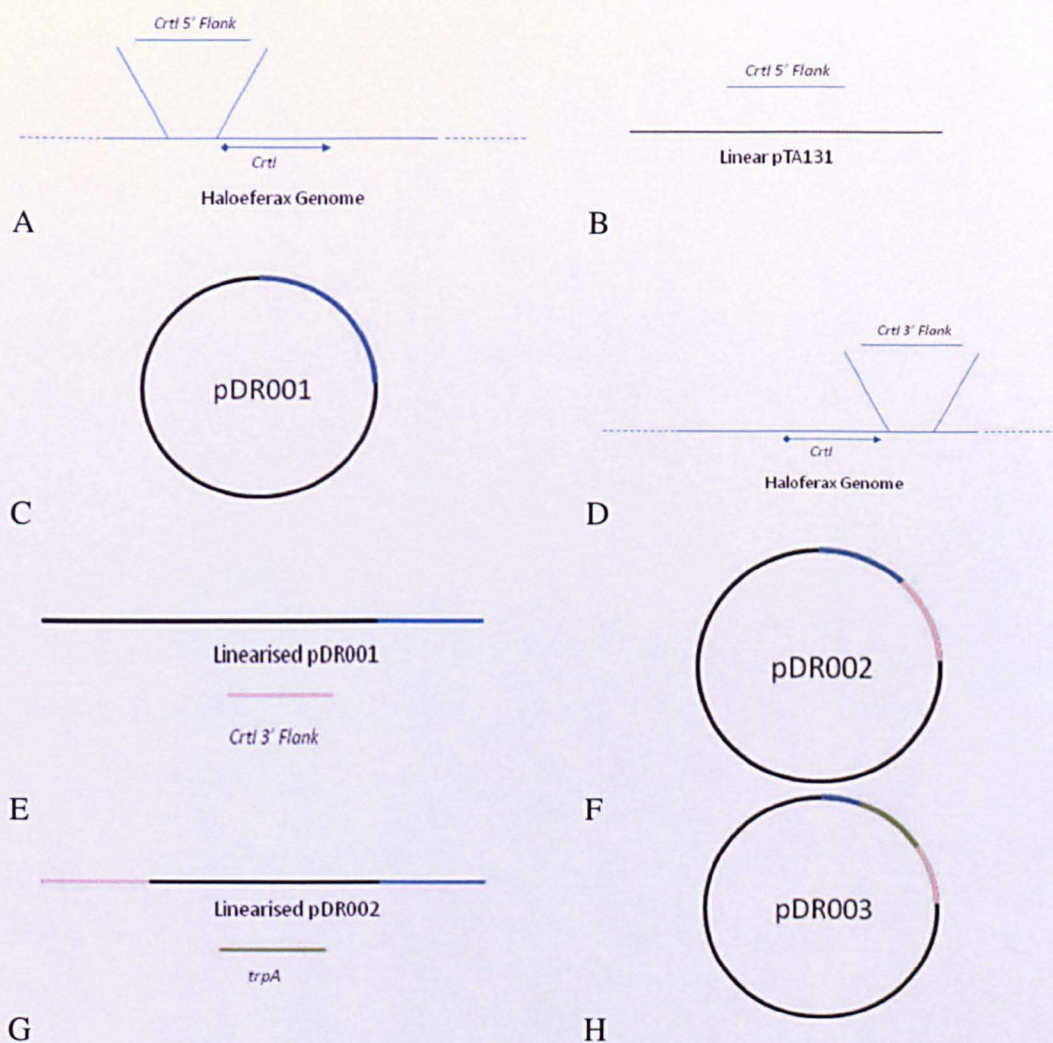


Fig 6.3 A schematic representation of the production of a knockout vector for *crtI*. **A** A PCR product homologous to the 5' flank region of *crtI* was produced. **B** 5' Flank and pTA131 plasmid were digested with KpnI and BamHI. **C** *crtI* 5' Flank was ligated into pTA131 giving pDR001. **D** A PCR product homologous to the 3' flank region of *crtI* was produced. **E** 3' Flank and pDR001 plasmid are digested with SpeI and BamHI. **F** *crtI* 3' Flank is ligated into pDR001 giving pDR002. **G** pDR002 and pTA298 are digested with BamHI. **H** TrpA is ligated into pDR002 giving pDR003.

accordance with previous bioinformatics results (Chapter 3).

## 6.4 Cloning of *crtI* for Heterologous Protein Expression

To further investigate the final enzyme in the lycopene biosynthesis pathway, *crtI*, attempts were made to clone and over express *crtI*. Primers were designed that would produce a PCR product containing the *crtI* gene by searching an electronic copy of the genome and selecting sections of DNA adjacent to the gene. The sites selected were chosen to allow for cloning into the pET28a system. For this reason the enzymes NdeI and XhoI were selected as the restriction sites to be inserted. NdeI has the restriction site GATATG and as such can be used to accurately ensure the position of the start codon in the plasmid construct after cloning with minimum effort. Once primers were designed (Table 2.2) it was next necessary to perform the PCR.

PCRs were run using the designed primer pair with genomic DNA extracted from *H. volcanii*. The PCR was run on a temperature gradient, in order to determine the best annealing temperature to use. The annealing temperature was selected (Table 2.2) and a DMSO gradient was used in an attempt to improve yield of PCR product without increasing mispriming as much as lowering the annealing temperature would.

The DMSO gradient showed high concentrations of DMSO (15%) gave the highest yields with no secondary products for this PCR. The PCR was scaled



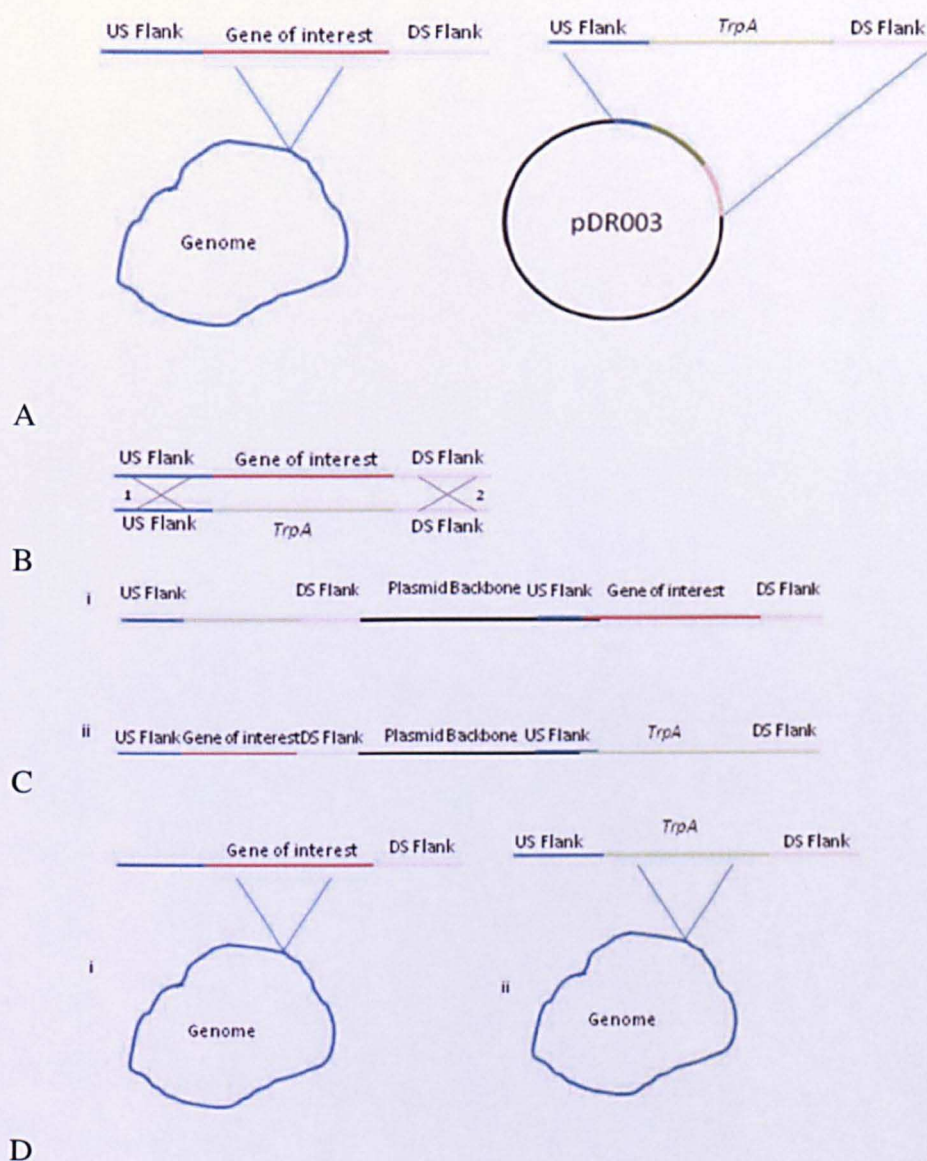


Fig 6.4 **A** Transformation introduces the plasmid into the cell, allowing it to interact with the genome. **B** Homologous regions of the plasmid and genome interact in recombination events at either region 1 or 2. **C** These events produce either i or ii. **D** A second crossover event, will produce one of two outcomes, with subsequent plasmid loss, dependant on whether the event happens at the same flank as the cross in event i or at the opposite flank ii.

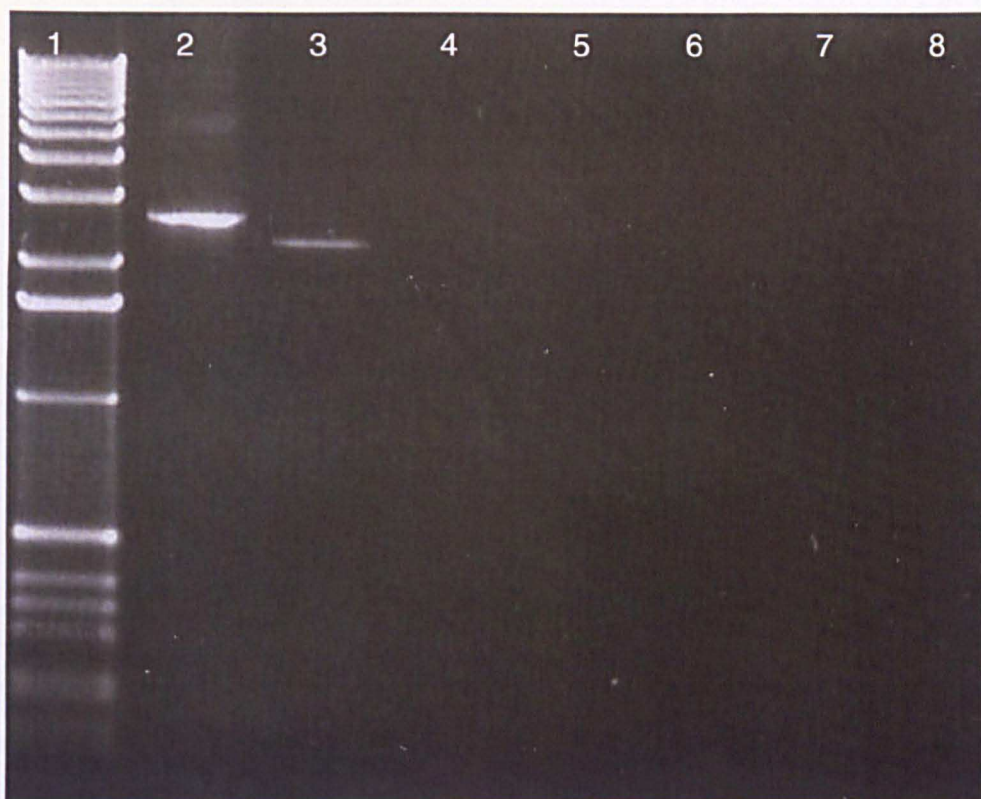


Fig 6.5. Confirmation of *crtI* knockout Lanes 1 shows promega X ladder, lane 2 shows a PCR product from H53 indicating the presence of *crtI*, lane 3 shows a PCR product from the mutant, lane 4 shows a PCR performed using only the upstream primer on H53 gDNA, lane 5 shows a PCR performed using only the downstream primer on H53 gDNA, lane 6 shows a PCR performed using only the upstream primer on mutant gDNA, lane 7 shows a PCR performed using only the downstream primer on mutant gDNA, Lane 8 is empty



CLUSTAL W (1.7) multiple sequence alignment

```

crtI- Predicted      GTTTCACGCTTGGAGTCCGCGATGTCGCCGTCGAACTCGGCGGGGTCGGCGTCGGTCCG
crtI- Sequence      -----
crtI- Predicted      CCGCCACAGCAACAGATCGGCCACGTCGTCGGGGTCGGGCCCTGCCCGCGGTGAGGT
crtI- Sequence      -----CGGTGAGGT
crtI- Predicted      CGGTGGCGACGAGGCCGGGTCGACGACCGTCACGGTCCGGTCGCAGTCGGCGGCGAACT
crtI- Sequence      CGGTGGCGACGAGGCCGGGTCGACGACCGTCACGGTCCGGTCGCAGTCGGCGGCGAACT
crtI- Predicted      GCCGGACGAGCGCTCGGCGGCCGCTTGGAGACGGCGTACGCGCCATCCCCGGTTTCG
crtI- Sequence      GCCGGACGAGCGCTCGGCGGCCGCTTGGAGACGGCGTACGCGCCATCCCCGGTTTCG
crtI- Predicted      CCTCGCGGCGATGGAGCCGAGGGGACGAGGATGCGCCGTCGTCGGCCATGTGCGGCA
crtI- Sequence      CCTCGCGGCGATGGAGCCGAGGGGACGAGGATGCGCCGTCGTCGGCCATGTGCGGCA
crtI- Predicted      GCGCCTCTTTGACCGCGCGAAGACGCGCGGACGTTGGTCTGAGGGTGTGTCGAACG
crtI- Sequence      GCGCCTCTTTGACCGCGCGAAGACGCGCGGACGTTGGTCTGAGGGTGTGTCGAACG
crtI- Predicted      CCGCGTACGAGTCTCGGCGCGGGCATCTCGCGGGCGTCCCGTGGCGACGGCGCGT
crtI- Sequence      CCGCGTACGAGTCTCGGCGCGGGCATCTCGCGGGCGTCCCGTGGCGACGGCGCGT
crtI- Predicted      TGGCGACGAGTACGTGATGCGTCCGCGGCGGGCGCGCTCTCCATCAGGCGCTCCA
crtI- Sequence      TGGCGACGAGTACGTGATGCGTCCGCGGCGGGCGCGCTCTCCATCAGGCGCTCCA
crtI- Predicted      TGTCGAGTTCTGTCGCGACGTCGGCGCGGACTGCGCTGGCGCTGCCCGCTCGCTTCA
crtI- Sequence      TGTCGAGTTCTGTCGCGACGTCGGCGCGGACTGCGCTGGCGCTGCCCGCTCGCTTCA
crtI- Predicted      TATCGTTGACGACCGACTGGAGGGCGTCGCGCTCGCGGGCGCAGGCGACGACCGTCGCG
crtI- Sequence      TATCGTTGACGACCGACTGGAGGGCGTCGCGCTCGCGGGCGCAGGCGACGACCGTCGCG
crtI- Predicted      CCGCGCGACCGAGGGCGCGCGACCGCGCGCGGATTCCGGAACCTCGCACCGGTGACCA
crtI- Sequence      CCGCGCGACCGAGGGCGCGCGACCGCGCGCGGATTCCGGAACCTCGCACCGGTGACCA
crtI- Predicted      CCGCGGTGGTGTGTTTCATGCGGGCGCGTAAGGAGGGATCCCGTGGATAAAACCCCTCGT
crtI- Sequence      CCGCGGTGGTGTGTTTCATGCGGGCGCGTAAGGAGGGATCCCGTGGATAAAACCCCTCGT
crtI- Predicted      TGACGCGCGGGCAGCGTCGTGATGGCGGTACACCGGGGCTTCGTGGCAGTACGCTGG
crtI- Sequence      TGACGCGCGGGCAGCGTCGTGATGGCGGTACACCGGGGCTTCGTGGCAGTACGCTGG
crtI- Predicted      CCCGAACAGCAACTACTATGCGTTCGGAAGCCGAACCTCTGCAGCCATGTGCTCGAAGAC
crtI- Sequence      CCCGAACAGCAACTACTATGCGTTCGGAAGCCGAACCTCTGCAGCCATGTGCTCGAAGAC
crtI- Predicted      GCCTTCTCCGACGGCCCGGGCTTCGTCCCTATCTTGGCGCGGGCAGCCCGACTACGAG
crtI- Sequence      GCCTTCTCCGACGGCCCGGGCTTCGTCCCTATCTTGGCGCGGGCAGCCCGACTACGAG
crtI- Predicted      TCCTCGCTCGAATACGTGAGGCGCTCGAAGCGGGCGCGGACGTCATCGAATCGGA
crtI- Sequence      TCCTCGCTCGAATACGTGAGGCGCTCGAAGCGGGCGCGGACGTCATCGAATCGGA
crtI- Predicted      CTTCCGTTCTCGGAGCCCATCGCCGAGGGGCGGACCATCCGAAACGCGGTTGTTCGGTCCG
crtI- Sequence      CTTCCGTTCTCGGAGCCCATCGCCGAGGGGCGGACCATCCGAAACGCGGTTGTTCGGTCCG
crtI- Predicted      CTCGAAGGCGGCATGACGCCGACGCGCTTTTCGAGTTCTGTCGAGGACCTCGACGTGTCG
crtI- Sequence      CTCGAAGGCGGCATGACGCCGACGCGCTTTTCGAGTTCTGTCGAGGACCTCGACGTGTCG
crtI- Predicted      GTGCCGCTGGTCTGTATGAGGTACTACAACCTCATCTACGTACGGCGATGAACCCGGA
crtI- Sequence      GTGCCGCTGGTCTGTATGAGGTACTACAACCTCATCTACGTACGGCGATGAACCCGGA
crtI- Predicted      CCGCGGCGGTCGTCGAGAAGGCGGCGGAAGTCGGTATCGAGGGCTTCGTGTCGCCGAC
crtI- Sequence      CCGCGGCGGTCGTCGAGAAGGCGGCGGAAGTCGGTATCGAGGGCTTCGTGTCGCCGAC
crtI- Predicted      CTGCGGCGGTCGAGGAGGCGGCGGCTCCGCGAGGCGCTGCGACGAGTTCGGCCTCGACCTC
crtI- Sequence      CTGCGGCGGTCGAGGAGGCGGCGGCTCCGCGAGGCGCTGCGACGAGTTCGGCCTCGACCTC
crtI- Predicted      GTGTTTCATCGTCGCGCCGACGACCCGCGGCGAGCGCTCGACCGAATCATGGAACAGGTC
crtI- Sequence      GTGTTTCATCGTCGCGCCGACGACCCGCGGCGAGCGCTCGACCGAATCATGGAACAGGTC
crtI- Predicted      TCGGGCTACGTCTACGTGACGCGCGCTCGGCACGACGGGCGCGAGTCGAGCGTCTCC
crtI- Sequence      TCGGGCTACGTCTACGTGACGCGCGCTCGGCACGACGGGCGCGAGTCGAGCGTCTCC
crtI- Predicted      GACCAGACCGACTCGTCGCTCGAAGCACTACCGACTACGACGTGCCAAGGCGGTCCGGC
crtI- Sequence      GACCAGACCGACTCGTCGCTCGAAGCACTACCGACTACGACGTGCCAAGGCGGTCCGGC
crtI- Predicted      TTCGGAATCAGCGACGGCGACACCGCGGAGCGCATCGTCGCGAGCGGTGCCGACGGCATC
crtI- Sequence      TTCGGAATCAGCGACGGCGACACCGCGGAGCGCATCGTCGCGAGCGGTGCCGACGGCATC
crtI- Predicted      ATCGTCGCGAGCGCGCTCGTCGACATCGTCGCGAGGGGCGGAGAACGCGGACGACGCC
crtI- Sequence      ATCGTCGCGAGCGCGCTCGTCGACATCGTCGCGAGGGGCGGAGAACGCGGACGACGCC
crtI- Predicted      GAAACGGTCGCGGACCGACTCGAAGCGCTCGCCCGGAACTCGAAGACGGTGCGGTAGCG
crtI- Sequence      GAAACGGTCGCGGACCGACTCGAAGCGCTCGCCCGGAACTCGAAGACGGTGCGGTAGCG
crtI- Predicted      GGCAGCTCGCAACGCCACCGCATCCGGAACGCACATAACGGATCCCGACTCGTCTACCT
crtI- Sequence      GGCAGCTCGCAACGCCACCGCATCCGGAACGCACATAACGGATCCCGACTCGTCTACCT
crtI- Predicted      CGCGGTCTCTCGCGGCCCGGTTCTGGCTCTACCTCGCGGGCCCGTCTGTCGCGCGT
crtI- Sequence      CGCGGTCTCTCGCGGCCCGGTTCTGGCTCTACCTCGCGGGCCCGTCTGTCGCGCGT
crtI- Predicted      CGCGGCCGCGGAGTGCCCTCGCCGACCTGTTCCGCTCGAACCCTGACCCCTGTTCCG
crtI- Sequence      CGCGGCCGCGGAGTGCCCTCGCCGACCTGTTCCGCTCGAACCCTGACCCCTGTTCCG
crtI- Predicted      GTACTTCTCGTCCCGCGAAGCTCTTCTCTACGGCGTCAACGACGTGTTTCGACGCCGA
crtI- Sequence      GTACTTCTCGTCCCGCGAAGCTCTTCTCTACGGCGTCAACGACGTGTTTCGACGCCGA
crtI- Predicted      CGTGGACGAAGCGAACCCGAAGAGGACGACCGCGAGGCGCGTGGCGCGGCGACCCCGT
crtI- Sequence      CGTGGACGAAGCGAACCCGAAGAGGACGACCGCGAGGCGCGTGGCGCGGCGACCCCGT
crtI- Predicted      GAAACACCGTCGTGTCGCGCGAGCGGCTGTCGGAGTCGGCCTGTTCCGCGTCTGCTCC
crtI- Sequence      GAAACACCGTCGTGTCGCGCGAGCGGCTGTCGGAGTCGGCCTGTTCCGCGTCTGCTCC
crtI- Predicted      GCGGGTGGCGTGGCGGTGGCTCGCCGCGCACTTCTTCTCGCGCTCGAGTACAGCGCGCC
crtI- Sequence      GCGGGTGGCGTGGCGGTGGCTCGCCGCGCACTTCTTCTCGCGCTCGAGTACAGCGCGCC
crtI- Predicted      GCCGTTTCGGTTCAAGACGACCGCGCTTCTCGACTCGGTCTCGAAGCGGCTGTACGTCTC
crtI- Sequence      GCCGTTTCGGTTCAAGACGACCGCGCTTCTCGACTCGGTCTCGAAGCGGCTGTACGTCTC
crtI- Predicted      GCGGGGCGTCCGCGCTACGCGGCGCTCTCGGGGTGAAACCCCGATGCTCG-----
crtI- Sequence      GCGGGGCGTCCGCGCTACGCGGCGCTCTCGGGGTGAAACCCCGATGCTCG-----
crtI- Predicted      CGACCGGAGGCGGTATTCGAACGACCGCGACG
crtI- Sequence      -----

```

Fig 6.6. Sequence alignment of predicted  $\Delta$ *crtI* sequence containing *trpA* (highlighted) with sequencing data.

up to a larger total volume (50  $\mu$ l). The scaled up PCR produced sufficient yield of PCR product which was then taken forward to cloning.

The PCR product and pET28a vector (Section 2.3) were digested in double digests using Fermentas FastDigest enzymes and buffer. The digested vectors were treated with Antarctic phosphatase and purified before ligation was performed (Section 2.16). Ligated plasmid was transformed into *E. coli* Top10 cells by electroporation (Section 2.17) before being plated on kanamycin containing media. The resulting transformation efficiency was calculated to be approximately 50000 cfu/ $\mu$ gml. Whilst this indicated the transformation had been successful it was then necessary to check that the correct plasmid had been constructed.

Several colonies were selected as candidates to be mini prepped and analysed to ensure the correct plasmid had been created. 83.3% of the selected colonies contained plasmid that gave the correct restriction profile, these plasmids were then sequenced to ensure that no undesired mutations had occurred. The sequencing results confirmed that the correct inserts with no undesired mutations had been ligated into the plasmid. The plasmid construct, pDR007 was then transformed by electroporation into BL21 cells for expression.

Expression was carried out in kanamycin containing LB broth in order to produce recombinant *crtI* protein. Induction was performed with 1 mM IPTG, as pET28a is lac suppressed. Samples were taken from the culture each hour for four hours after induction and the whole culture collected after four hours

by centrifugation. The samples were then run on an SDS PAGE gel in order to discover if expression had occurred successfully. No expression was seen across four hours (Fig 6.7) as such it was concluded that *crtI* may be unable to express in this host, possibly due to codon incompatibility or due to the cytosolic conditions of *E. coli* being inappropriate for correct folding of the protein. Repeating the experiment and transforming the plasmid into the *E. coli* BL21 Rossetta cell line could alleviate the problem of codon usage, as the Rossetta cell line contains a plasmid encoding rare tRNAs.

pDR007 was transformed successfully into *E. coli* BL21 Rossettas. The expression experiment was then repeated to discover if the codon usage problem was in fact the cause of the failed expression. No expression was seen at any point, it was therefore concluded that expressing *crtI* in *E. coli* was unlikely to work, possibly due to incompatible cytosolic conditions. At this point attempts to express the protein were abandoned.

## 6.5 Discussion

The strains containing UV induced mutations did not contain the disruptions in the carotenoid synthesis pathway sufficient to abolish lycopene synthesis. The reason for the apparently white phenotype of these strains when grown on plates is therefore not the deletion of the carotenoid pathway, rather it may be due to a down regulation of the pathway. What mutations may cause a down regulation of the carotenoid pathway is a matter of speculation as a large array



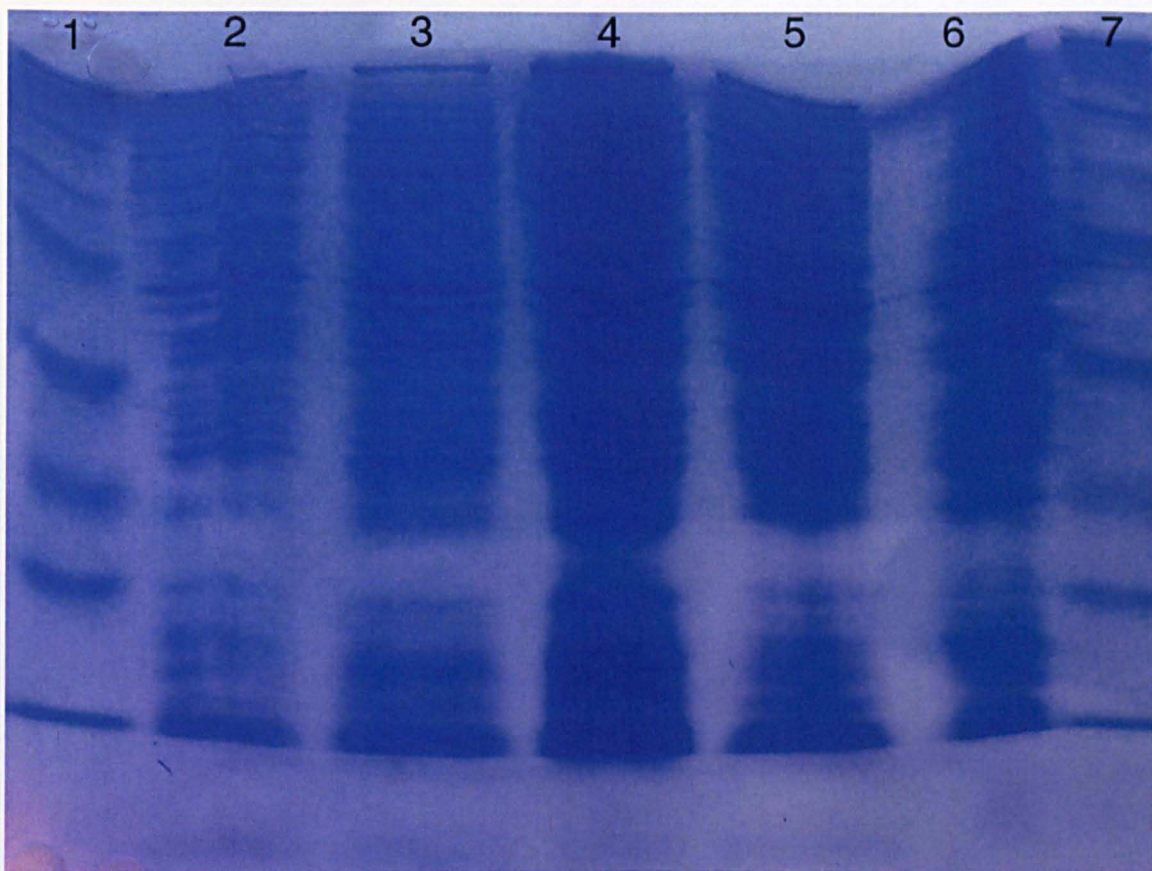


Fig 6.7 Expression of *crtI* in BI21 Star Rosettas. Lanes 1 and 7 are Promega Broadrange Molecular Weight Marker, Lanes 2-6 show whole cell samples taken at 0, 1, 2, 3 and 4 hours respectively.

of possibilities exist. A mutation in one of the genes in the pathway itself that does not eliminate, but attenuates functionality is certainly a possibility, however it is equally possible that such a mutation could have occurred earlier in the isoprenoid pathway. It is also feasible that it is not one of the genes directly responsible for the production of carotene or one of its precursors that has produced this phenotype, but rather a mutation in the genes of one of their regulating factors, or indeed in genomic elements recognised by such factors. As such it is difficult to speculate on what the meaning of any observed behaviours of these organisms might be. For this reason it seems prudent to disregard the UV induced mutants as there is little of value that can be derived from them.

Accordingly the targeted gene knockouts take on greater import as they are now the only available method for studying *H. volcanii* without lycopene. The approach used to produce these knockouts utilises strong selection against undesired strains at several stages, in addition to taking advantage of *H. volcanii*'s tendency to freely undergo homologous recombination. The *crtI* knockout was focussed upon as the most crucial, as it would remove lycopene with the least disruption to the organism. Initially it seemed that the *crtI* knockout, like the *pmdc* knockout failed to grow after transformation, which would indicate that the gene may be essential. After a long period of incubation at room temperature, however it was seen that the *crtI* knockout did in fact grow, simply very slowly. This slow growth shows that, whilst lycopene production is not essential, it is important for the maintenance of

healthy cells. It is possible that the slow growth was caused by a protective effect of lycopene no longer being present.

It is interesting to note the failure of the strain containing the *pmdc* deletion to grow. This is likely to have been caused by deletion mutants being non viable, rather than the transformation having failed, as control transformations succeeded. Whilst this is not definite, it is a good indication the *pmdc* is essential, which is unsurprising given its position in the biosynthetic pathway (Fig 3.1) producing either a precursor to isopentyl pyrophosphate, isopentyl phosphate, or isopentyl pyrophosphate itself and thus leading to the entire isoprenoid pathway. If *pmdc* is indeed essential it confirms that only the route from acetyl coA and not the route from pyruvate to isopentyl pyrophosphate is present in *H. volcanii*. This could be tested simply by the addition of a recovery plasmid containing a cloned *pmdc* to a suitable strain (such as H53) before transformation with the knockout vector. This would ensure continued survival of the cell and allow confirmation of the knockout prior to curing of the plasmid.

The attempt to express cloned *crtI* failed, despite the initial success of the cloning and prevented investigation of the protein. This is likely due to a discrepancy between *E. coli* and *H. volcanii* in the processing of genetic information or proteins. As rare tRNAs were provided, the failure to express was likely caused by a failure in post translational modification or protein folding caused by an incompatibility between the transgenic protein and its

host. The cause of this could feasibly be attributed to misfolding of the protein in what is, for a halophilic protein, a low salt cytosolic environment in *E. coli*. Such a misfolding would likely see the protein rapidly degraded by *E. coli* proteases and as such the lack of apparent expression. It is also possible that the failure was indeed translational, however this seems less likely due to the compensatory nature of the rare tRNAs included.

The products of the molecular alterations described in this chapter have been carried forward to allow examination of the effects of lycopene removal on *H. volcanii*, the results of which are discussed in Chapter 7.

## **Chapter 7**

# **The Carotenoid Metabolite Profile of *Haloferax volcanii* Mutants and Their Response to Stress Conditions**



## **Chapter 7 The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions**

### **7.1 Introduction**

The understanding of the function of a compound in an organism can often best be shown by removing it. This may be accomplished by removal of the gene responsible for producing the compound. Doing so allows comparative studies to be performed between the wild type organism possessing the compound and the mutant, without it. Differences between the wild type strain and that incapable of producing the compound of interest are then attributed to that compounds lack.

This approach allows for both general questions regarding the compounds role in normal growth and specific questions about its function a variety of narrow conditions to be addressed. The most obvious question to answer is whether or not a compound is essential to the cell, if the cell dies with its removal this shows it was essential, if it continues to survive it shows it was not. Following from this, even in the case of non-essential compounds a group of questions relating to the impact of the health of the mutant can shed light on its importance. Assuming the mutants are viable it is then possible to interrogate them under carefully controlled conditions to ascertain if specific

Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions  
environmental factors affect the health of the organism deficient in the compound of interest.

These techniques are not without limitations, however, as the removal of a gene may have effects beyond the removal of the product of an enzyme for which it encodes. Confounding factors include the possibility that the region of DNA codes for multiple genes, either by splicing, common in eukaryotes, by multiple genes overlapping, often on opposite strands, more common in prokaryotes and viruses, or even by frame shifted or truncated versions of the gene also coding for a viable protein. Beyond these considerations if the gene removed codes for a protein that is part of a pathway then more than simply its product will be removed from the cell. Similarly if an enzyme has multiple products, or if a feedback mechanism within the cell responds to the lack of the gene, the enzyme or the product other factors may account for the observed changes.

Despite these limitations and provided they are kept in mind when analysing results from these techniques the data acquired can provide some of the strongest evidence for function available.

## **7.2 Lycopene Deficient Mutant *H. volcanii* Growth**

The behaviour of mutants compared to wild type organisms under normal growth conditions can tell us about the effects of the induced mutation. In this

## Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions

case the mutant *H. volcanii* strain  $\Delta crtI$  (Section 2.1) was grown under the same conditions as *H. volcanii* DS70 and their behaviours compared. On plates it was found that the media used caused a greater difference in growth time to be evident. On HvYPC plates  $\Delta crtI$  could be grown to pickable colonies within 2 months, approximately 8 times as long as the wildtype strain. On chemically defined media (Hv NA-Lac, Section 2.8) growth to this level was only achieved once and took over 8 months, more than 16 times as long as wildtype. From this we can see that lack of the *crtI* gene either slows growth or increases killing significantly under these culture conditions. To see if this was particular to growth on plates, or a general effect, comparisons were performed in liquid media.

*H. volcanii* is an aerobic organism and as such oxygen must be supplied throughout growth. Under laboratory conditions in batch culture it is therefore impossible to prevent some evaporation occurring as flasks cannot be completely sealed. Consequently there is a limit to the amount of time a culture can be grown under these conditions. The addition of fresh media would alter the culture significantly enough, due to increased nutrients and dilution of waste products, to cause the experiment to no longer be considered as a single growth. Similarly the addition of water to compensate for that which had evaporated would cause localised osmotic shock, killing organisms and once more making it impossible to compare the experiment to one performed in a single growth. Growth under liquid culture conditions was

Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions  
therefore stopped after approximately 10% of the volume of the culture had evaporated (3-6 days).

Within the timeframe accorded by the above restrictions no growth of  $\Delta crtI$  was seen in Hv Na-lac media (Section 2.8) under standard growth conditions (Section 2.9) with 50 ml media in a 500 ml flask. By contrast the wildtype grew normally. This suggested that the organism either could not grow in liquid culture or that some of the growth conditions prevented it from doing so. To investigate this it was decided to attempt to grow the organism in HvYPC (Section 2.7).

HvYPC is a rich medium, known to be more conducive to *H. volcanii* growth than chemically defined media such as Hv Na-Lac. Unfortunately because it is not a chemically defined medium it is impossible to be precise when describing its formulation as yeast extract and peptone both have varying compositions. Similarly to the previous experiment 50 ml volumes were used in 500 ml flasks. Again no growth was seen in  $\Delta crtI$ . Either growth in liquid culture was not possible for the strain or some other condition needed changing. As it was known that exposure to reactive oxygen species (ROS) could cause an increase in lycopene production in the wildtype strain (Fig 5.9.) it was hypothesised that reduction in oxygen levels may have a protective effect on the strain.

## Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions

To test this a range of culture volume to flask volume ratios were used, from 10% to 70%. To allow a larger number of ratios to be studied simultaneously flask volumes were decreased to 50 ml. HvYPC was still used as media to maximise the chances of growth. It was found that under these conditions that growth appeared in flasks between 25% and 35% flask volumes to detectable levels within the period of the experiment. Growth was not seen in the wildtype strain above 50% flask volume (Table 7.1). These results suggest that whilst a higher ratio of culture volume to flask volume does give some protective effect to the mutant strain the resultant lack of oxygen may itself retard their growth. It was decided that scaling up the successful growths in order to allow sufficient culture to be collected for an extraction of carotenoids was an appropriate next step.

To allow sufficient  $\Delta crtI$  to be harvested so that its carotenoids could be extracted culture sizes were scaled up. Starter cultures were grown in 50 ml flasks using a 30% culture volume before being transferred to 500 ml flasks again with 30% culture volume. This, however resulted in no growth in the larger flask, even with high inoculums. This suggested that the flask size as well as culture volume to flask volume ratio was having an effect on culture growth. This could be due to more efficient mixing of oxygen into the culture in larger flasks under the same rotational speeds. It was therefore suggested that a smaller flask may prove more beneficial.

Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions

Table 7.1 Optical densities reached by *H. volcanii*  $\Delta crtI$  and *H. volcanii* DS70 grown in HvYPC under varying culture to flask volume ratios. All samples are means taken from three replicates and variation shown is standard error

Culture to Flask Volume Ratio (%)	<i>H. volcanii</i> $\Delta crtI$ Growth (OD <sub>600nm</sub> )	<i>H. volcanii</i> DS70 Growth (OD <sub>600nm</sub> )
10	0.0	>1.0
15	0.0	>1.0
20	0.0	>1.0
25	0.21 $\pm$ 0.01	>1.0
30	0.28 $\pm$ 0.03	>1.0
35	0.19 $\pm$ 0.02	>1.0
40	0.0	>1.0
45	0.0	>1.0
50	0.0	0.53 $\pm$ 0.04
55	0.0	0.0
60	0.0	0.0
65	0.0	0.0
70	0.0	0.0
75	0.0	0.0

## Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions

250 ml flasks were employed using 35% culture to flask ratios and some growth was observed, although it appeared slower than that in the smaller cultures. It was decided that as growth was being detected a resumption of the use of Hv Na-Lac should be attempted in order to allow experiments with hydrogen peroxide shock, similar to those showing lycopene build up in the wildtype (Fig 5.8).

Growth in Hv Na-Lac was attempted at between 25% and 35% culture volume to flask volume ratios in both 50 ml and 250 ml flasks. Growth was detected only in 50 ml flasks and was very slow, reaching OD<sub>600nm</sub> 0.1 AU within the time frame of the experiment (3-6 days) and showing little subsequent increase in density even when allowed to grow for longer (1-2 weeks). It was therefore determined that experiments requiring Hv Na-Lac would have to be performed in small scale cultures.

Throughout the above experiments no  $\Delta crtI$  culture exceeded OD<sub>600nm</sub> 0.3 AU and whilst their behaviour post lag phase mirrored that of the wildtype they quickly ceased growing again. This may have been due to the evaporation of water causing the concentrations of media components to become non supportive of further growth, however rather than an inherent inability of the organism to sustain further growth.

### 7.3 $\Delta crtI$ Carotenoid Metabolite Profile

It was vital to establish that the deletion of *crtI* from the genome of the organism had indeed removed its capacity for producing lycopene. In addition the question as to whether it would still produce bacterioruberin and how the phytoene levels would respond to the absence of lycopene were interesting questions.

Four 250 ml flasks were prepared in which  $\Delta crtI$  was grown to a density of OD<sub>600nm</sub> 0.2 AU. These cultures were harvested and from them carotenoids were extracted in the usual way (Section 2.10-2.11). This was run using the gradient HPLC method (Section 2.12) and a chromatogram produced (Fig 7.1). As can be seen on the chromatogram only two peaks are present at 450 nm. The spectra of these shows in both cases a single peak absorbance at less than 300 nm with a gradual reduction in absorbance resulting in the reading at 450 nm. These peaks were not noted in the wildtype chromatogram, however as they are of low intensity comparative to the carotenoids, lycopene and  $\beta$ -carotene, and at similar retention times to those carotenoids it is likely that they were masked in these chromatograms. All carotenoids normally seen (lycopene, phytoene and bacterioruberin) are absent from this chromatogram. It is important to note that as all of these compounds and the unidentified carotenoids, presumed to be bacterioruberin intermediates, have been eliminated from the cell it is possible that the lack of any of these compounds



Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions  
may be the cause of the results seen in the mutants behaviour, rather than simply the lack of lycopene.

#### **7.4 $\Delta crtI$ Survival Under Stress Conditions**

As the mutant contains no detectable carotenoids attempting to study the behaviour of its carotenoid profile in response to stress in the manner applied with the wildtype (Chapter 5), would be fruitless. The response of the mutant to stress stimuli would therefore be measured in a different way, allowing direct comparison of its fitness to that of the wildtype under varying stress conditions.

Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions

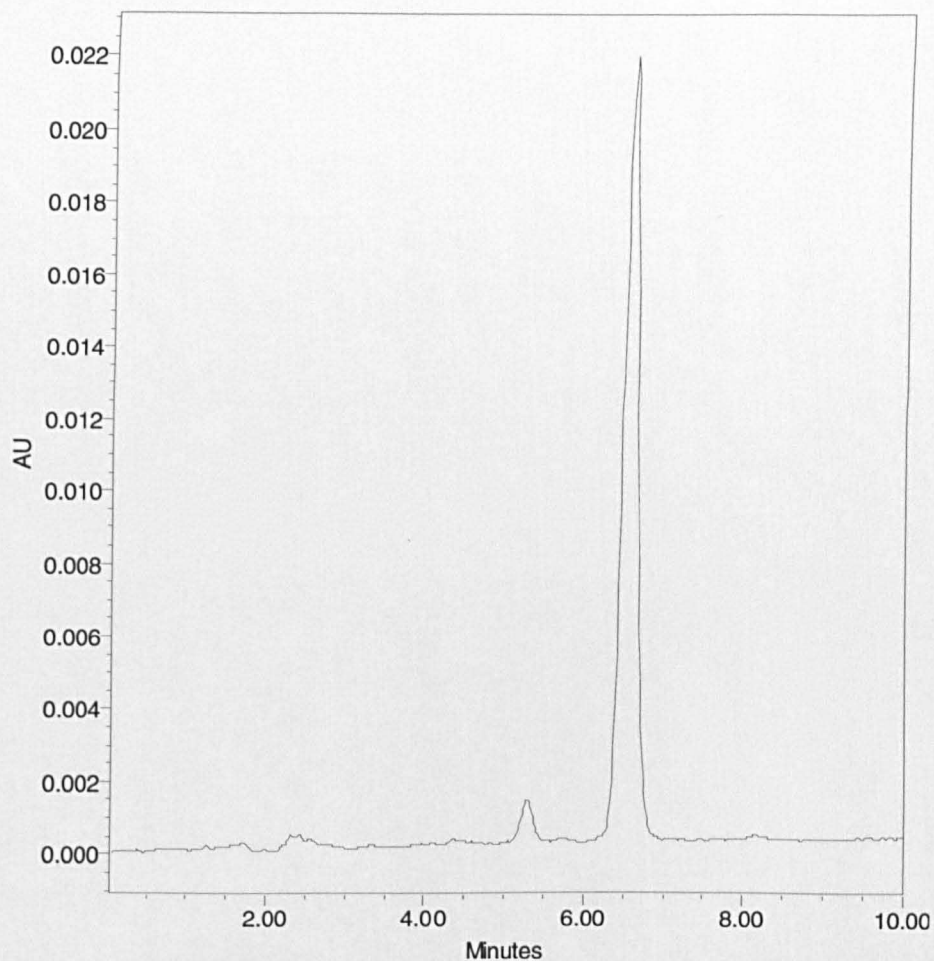


Fig 7.1. Chromatogram of extract from *H. volcanii*  $\Delta crtI$  culture, showing two unidentified peaks (5.25 mins and 6.5 mins). Absent are peaks for lycopene, phytoene and bacterioruberin.

## Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions

Resistance levels to various stressors can be measured by the surviving fraction of cells at a steadily increasing concentration of the stress. The resistance levels of  $\Delta crtI$  and *H. volcanii* DS70 were compared in this way, measuring their response to H<sub>2</sub>O<sub>2</sub> and to osmotic shock. In response to osmotic shock, over a range of salt concentrations from 20% to 3% both organisms performed identically within error (Fig 2a). This indicates that the carotenoids do not affect the organisms ability to deal with this stress.

In contrast when exposed to H<sub>2</sub>O<sub>2</sub> shock from 0 to 50 mM (Fig 2b) the wildtype showed a slow decline, maintaining 60% viability at 50 mM whilst the mutant strain shows very poor resistance, decreasing to an 85% survival at 10 mM and having lost all detectable viability at 30 mM and higher concentrations. This result indicates that the removal of the carotenoids has significantly impacted the organisms ability to resist oxidative damage.

## 7.5 Discussion

The reduction of phytoene in the mutant strain cannot be a direct result of the removal of *crtI*'s catalytic activity as it is the precursor to lycopene, not its product. For this reason it must be assumed that a regulatory mechanism detecting either the lack of lycopene, bacterioruberin or both has acted to curtail the production of phytoene or that there is an interaction between the phytoene desaturase and phytoene synthase genes or proteins directly. As there

Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions

are no known non-carotenoid products of phytoene it is safe to assume that in the absence of *crtI* the production of phytoene ceases. It is likely that

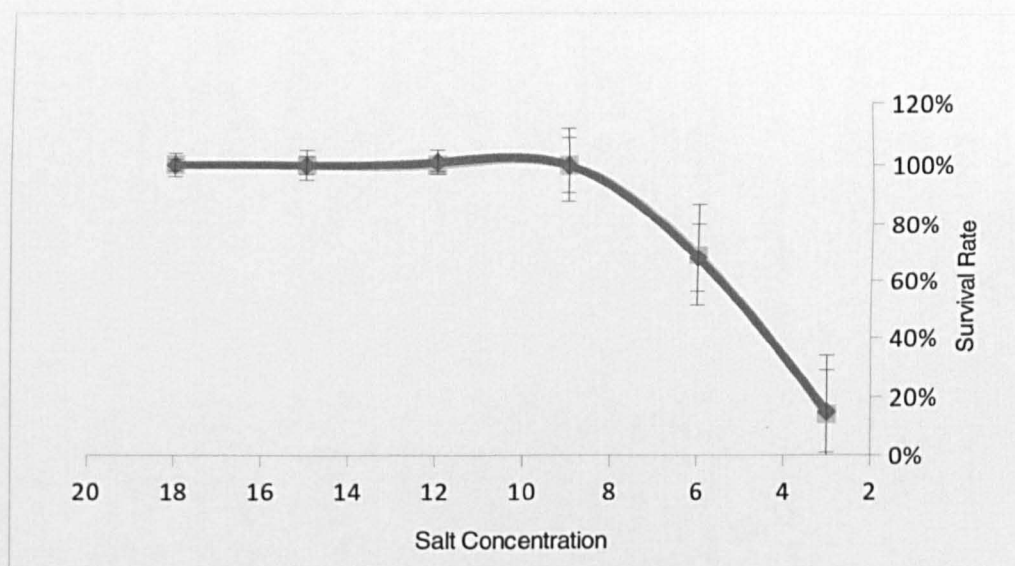


Fig 7.2a, Survival rates represented as a percentage surviving, 100% being taken as the survival in unstressed cells counted in the same way. Survival rate is plotted against salinity, where percentage represents the weight by volume percent of salt in the solution. 4 independent replicates make up each data point and standard error is represented by the error bars. Circular markers indicate wildtype, square markers mutant.

Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions

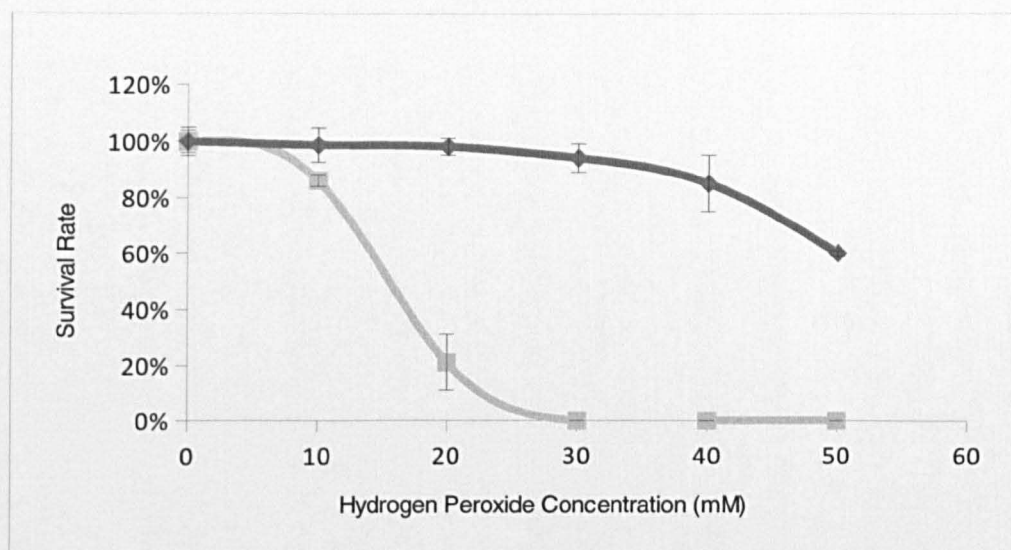


Fig 7.2b, Survival rates represented as a percentage surviving, 100% being taken as the survival in unstressed cells counted in the same way. Survival rate is plotted against H<sub>2</sub>O<sub>2</sub> concentration.. 4 independent replicates make up each data point. Circular markers indicate wild type, square markers mutant.

## Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions

compounds that would formerly have gone to make phytoene have instead been channelled into other products of the isoprenoid pathway.

The removal of bacterioruberin from the cell is most likely a direct result of the deletion of *crtI* as bacterioruberin is most commonly produced from lycopene (Britton *et al.* 1998). The removal of the other carotenoids suspected of being bacterioruberin intermediates indicates that they too are disrupted by this deletion which supports the suggestion that they are intermediates between lycopene and bacterioruberin, although it does not prove this.

The failure of  $\Delta crtI$  to grow beyond OD<sub>600nm</sub> 0.3 could be explained by its slow growth resulting in sufficient evaporation occurring over time to cause the media to become unable to support further growth of the mutant. This slow growth is itself indicative that even under normal culture conditions, without added stress the lack of carotenoids poses a severe challenge to *H. volcanii*, demonstrating that the carotenoids are an important, although not essential, product for cell viability.

The fact that the rate of growth increased in cultures with lower oxygen availability suggests that atmospheric levels of reactive oxygen species are able to overwhelm the other defence systems of *H. volcanii* (primarily DNA repair enzymes such as recA/B). This then would explain the maintenance of the carotenoid biosynthesis pathway despite its high energy cost in the wild

Chapter 7: The Carotenoid Metabolite Profile of *H. volcanii* Mutants and Their Response to Stress Conditions  
population, a population subjected to far higher levels of ROS than found in a laboratory setting.

This evidence is supported by the survival rates of wildtype *H. volcanii* as compared to the mutant when subjected to direct oxidative stress. Whilst the wildtype strain displayed a resistance similar to that of *D. radiodurans* (Tian *et al.* 2007) known to be a highly durable organism, the mutant displayed a great susceptibility to exposure to this stress. In contrast there is no difference in survival rate between the wildtype strain and the mutant when exposed to an osmotic stress, suggesting that the carotenoids provide a defence solely against oxidative damage and not as part of a general stress response.

# **Chapter 8**

## **Discussion**



## **Chapter 8 Conclusions and Discussion**

### **8.1 Major findings of this study**

The production of carotenoids in *H. volcanii* is important to the organisms growth and survival. Whilst initially this importance was ascribed to several different hypotheses (Chapter 1) it is possible now to state with some certainty that the carotenoids present in *H. volcanii* provide it with protection from oxidative damage (Chapter 7). *H. volcanii* produces three identified carotenoids, phytoene, lycopene and bacteriorhodopsin (Chapter 5) in addition to 3 unidentified carotenoids, which may be intermediaries between lycopene and bacteriorhodopsin (Chapter 5), which of these provides the protective effect is uncertain (Chapter 7).

### **8.2 Carotenoid biosynthesis in *H. volcanii***

Carotenoid synthesis in *H. volcanii* branches from the isoprenoid pathway (Chapter 3). The biosynthetic route to isopentyl pyrophosphate (IPP), the common precursor to all isoprenoids, has two potential forms. The first, common in bacteria, chloroplasts and green algae proceeds from pyruvate (Disch *et al.* 1998). The second, seen in all domains, starts at acetyl co-a and proceeds through mevalonate, (Disch *et al.* 1998). It is common among

## Chapter 8: Conclusions and Discussion

bacteria and some archaea for both pathways to be present, however only the mevalonate pathway was evident in *H. volcanii* (Fig 3.1).

Despite the fact that all the enzymes in this pathway are well characterised in other organisms (Chapter 3, [expasy.org/sprot/](http://expasy.org/sprot/)) one enzyme, essential to the pathway, could not be identified through the use of BLAST searches. Additionally two enzymes closely matched more than one BLAST hit and therefore their identity remains ambiguous.

These latter enzymes were initially identified as mevalonate kinase (Mvk) and mevalonate pyrophosphate decarboxylase (pMdc). Mvk showed close homology to both mevalonate kinase and mevalonate phosphokinase from other organisms (Chapter 3). pMdc showed equal homology to mevalonate phosphate decarboxylase and mevalonate pyrophosphate decarboxylase. The ambiguity as to the nature of these enzymes means that which enzyme in the pathway remains unidentified is unclear. The pathway either produces mevalonate pyrophosphate or isopentyl phosphate, therefore the unidentified enzyme is either a mevalonate kinase, mevalonate phosphokinase or isopentyl phosphokinase. A fourth possibility exists, if Mvk has function as both a mevalonate kinase and a mevalonate phosphokinase the pathway would be complete without the need to identify a further enzyme (Fig 3.1). This would mean that Mvk in *H. volcanii* displayed a unique functionality, not seen in other organisms

## Chapter 8: Conclusions and Discussion

Additionally there is at least one enzyme at the end of the carotenoid biosynthesis pathway, after formation of IPP, which was not elucidated by the BLAST searches performed, the enzyme or enzymes leading to bacterioruberin synthesis (Chapter 5). This pathway could lead either from lycopene or directly from phytoene, although the former is more common (Michal 1999) and found in related organisms (Siero C *et al.* 2003). The presence of a bacterioruberin in the cell shows conclusively that the synthesis pathway must be present, this is also true for the other unidentified carotenoids, although it is likely that they are bacterioruberin intermediates.

The elucidated carotenoid pathway then consists of a GGPP synthase, *IdsA*, a phytoene synthase, *crtb2* and a phytoene desaturase, *crtI*. This pathway also shows the high metabolic cost of carotenoid production to. The NADH and ATP cost of lycopene production is high enough that it in itself suggest that this pathway is not a conserved remnant. Additionally we know that the organism produces bacterioruberin and whilst the cost of the production of one molecule is unknown it is likely to be several times that of lycopene based on structure alone.

A high level of carotenoid production in stationary phase has been observed in other organisms (Heber *et al.* 2004). In the case of mycobacterium this has been shown to be regulated by SigF, an important sigma factor for regulating stationary phase growth and response to stress conditions. It is suggested that this may be due to the need to protect against internally synthesised reactive

## Chapter 8: Conclusions and Discussion

oxygen. As this study shows lycopene is produced at increased levels in stationary phase growth it seems reasonable to suggest this as the probable reason for such regulation in *H. volcanii* as well.

### 8.3 Model Proteins

The proteins *crtb2* and *crtI* are central to the carotenoid synthesis pathway. For this reason these proteins were modelled using the I-Tasser and Lommetts systems. These models give several interesting insights into the nature of these proteins.

*crtI* has a potential group of transmembrane helices (Chapter 3) that could anchor it firmly to the membrane. Although the prediction of these helices is not conclusive evidence of their existence, *crtI* being membrane associated is reasonable, particularly as both its substrate and product are to be found in the membrane (Palloza *et al.* 2010). This is not unprecedented in carotenogenic proteins, those found in several plants also have evidence of being membrane bound (Grunewald 2001). It is interesting then to see that the enzyme CrtB2 has no evidence of transmembrane helices. In fact, there is only one small patch on its surface which appears to be hydrophobic (Chapter 3), this consists of a short alpha helix, with only a slight transmembrane signature.

### 8.4 Deletion mutant

Whilst the high cost of the pathway indicates that it is likely functional other techniques are required to analyse that function. The removal of the pathway, or part thereof from the organism can give valuable information when answering this question. For this reason the production of mutant *H. volcanii* with genes in the carotenoid synthesis pathway removed was conducted. Two genes were successfully knocked out during the course of this study, *pmdc* and *crtI*, coding for the presumptive mevalonate decarboxylase and the phytoene dehydrogenase respectively (Chapter 6).

Knocking out the gene *pmdc* resulted in the organism failing to grow. Whilst this does not reveal anything about the function of the carotenogenic pathway it does allow some conclusions to be drawn about the more general isoprenoid pathway. pMdc is required for cell viability, which is unsurprising as it is a vital step in this pathway. The removal of pMdc is likely to lead to cell death by the removal of the mevalonate based isoprenoid pathway, indicating that the pyruvate pathway to isopentyl pyrophosphate is not present in *H. volcanii* and supporting the pathway developed by bioinformatics (Chapter 3).

The *crtI* knockout by contrast is viable, although its growth rate is greatly reduced (Chapter 7). This indicates that the *crtI* gene, whilst not vital for cell viability, is important in supporting normal cell growth. Given that the function of *crtI* according to our model is lycopene production it follows then

## Chapter 8: Conclusions and Discussion

that lycopene, or its products, play a key role in keeping the cell healthy. It is important to remember, however that the model pathway based on BLAST searches is not confirmed and therefore it was essential to check the carotenoid profile of *H. volcanii* to ensure it had been disrupted in the anticipated way. A carotenoid extract from *H. volcanii*  $\Delta crtI$  showed the presence of no carotenoids, not only was lycopene removed, but phytoene and bacterioruberin were also gone (Chapter 7).

The lack of all carotenoids from the cell means that all results gathered in Chapter 7 could be due to any of the carotenoids, or all of them, rather than to lycopene specifically. It also raises several questions, why were all carotenoids lost, including phytoene, lycopene's precursor and what caused the loss of these carotenoids. It is possible that the enzyme removed by the knockout, *crtI* is, rather than the phytoene dehydrogenase, the phytoene synthase. This is extremely unlikely however as it has a very strong homology to the phytoene dehydrogenases of related organisms (Chapter 3), whilst no homology to phytoene synthases. Another explanation must therefore be sought.

It is possible that the removal of *crtI* leading to the loss of lycopene and, presumably bacterioruberin, assuming bacterioruberin is synthesised from lycopene in *H. volcanii*, could produce a feedback loop preventing the production of phytoene and diverting resources elsewhere in the cell. Such a mechanism, however does not seem likely due to the fact that it responds to a condition that does not occur in the natural life cycle of the organism, a lack of

## Chapter 8: Conclusions and Discussion

all post phytoene carotenoids (Chapter 5). Further it would contradict the evidence showing an increase in phytoene when lycopene is severely depleted by oxidative attack (Chapter 5).

There is another possibility that presents itself however, it is reasonable to think that CrtI and CrtB2 may form a protein complex, as some plant carotenogenic enzymes do (Lopez 2008). If *crtI* is, as analysis suggests, anchored to the membrane by transmembrane alpha helices it is possible that CrtB2 then complexes to it, perhaps through its own short, hydrophobic alpha helix found on the exterior of the enzyme according to the model (Chapter 3). If this were the case it is feasible that removal of the *crtI* enzyme could lead to abolishment of phytoene synthesis. This could occur due to a failure on the part of the CrtB2 enzyme to take the correct conformation when not complexed to *crtI*, or simply be due to incorrect localisation.

Regardless of the specific mechanism it is clear that the removal of *crtI* blocks all carotenogenesis in *H. volcanii*. This has several effects on the cell, altering its growth rate and resistance to oxidative stress (Chapter 7). Although it is tempting to conclude that this means that lycopene acts to encourage growth and protect against oxidative damage, as hypothesised, it is not an accurate interpretation of the data. What can be concluded is that the carotenoids in *H. volcanii* offer it significant protection against hydrogen peroxide attack and, likely, against attack by other external ROS. It is not unreasonable to think that

it may also provide a protective effect from ROS generated within the cell, such as those produced by Fenton like reactions of several metals.

## 8.5 Response to oxidative stress

In its natural environment of the Dead Sea such stresses are likely to be encountered regularly. The dead sea has a high incidence of ultra violet radiation (Kudish 1998) and heavy metals leached from the surrounding area (Stiller 1979). Both UV radiation and heavy metals produce reactive oxygen species, which could then attack the cell. Without the carotenoids *H. volcanii* would not be able to protect itself from this damage and so it is safe to conclude that the carotenoid pathway is vital to the organisms survival in the wild.

The case for this is made stronger by the response *H. volcanii*'s carotenoid pathway shows to oxidative stress, with upregulation of phytoene production being seen in response to low levels of stress and upregulation of both phytoene and lycopene production seen in response to higher levels of oxidative stress (Chapter 5).

This upregulation effect may be due to the organism sensing that it is under oxidative attack and responding by increasing its protection. The pattern of initial depletion followed by subsequent replenishment seen in Fig 5.9,



## Chapter 8: Conclusions and Discussion

however suggests that this effect may be in response to the cell detecting depletion of the compounds beyond their standard levels.

### 8.6 Further study

In order to carry this study further the production of functional CrtI and CrtB2 proteins for in vitro analysis would be useful. This would allow localisation studies to be performed that may give an insight into whether these enzymes do form a complex. Additionally comparison of the effects of oxidative damage from other sources on the wildtype and  $\Delta crtI$  mutant would indicate if the protective effect against ROS is functional against all such stressors. Further, the response of the wildtype's carotenoid synthesis pathway to these stressors, including the response of bacterioruberin could shed further light on the subject.

It would also be of interest to clone and overexpress *mvk* in order to analyse its function. This would determine which route the biosynthesis pathway takes to IPP and confirm whether or not Mvk acts at multiple steps in the pathway.

Additional work could also be performed on *pmdc*, by attempting to obtain a conditionally lethal knockout of this gene. This could be accomplished by cloning *pmdc* into a vector containing an inducible, or repressible, promoter for use in *H. volcanii*. It should then be possible to make a knock out of this gene by inducing, or not repressing, said promoter throughout the knockout

## Chapter 8: Conclusions and Discussion

process, provided that process is performed on a strain transformed with the appropriate plasmid. Alternatively a copy of *pmdc* under control of such a promoter could be incorporated into the genome, either as part of the knockout process or via a second suicide vector in an alternate area of the chromosome.

Using either of the above methods could lead to the production of a  $\Delta pmdc$  mutant, which could then have induction removed or repression applied to investigate the effect of reduced or ablated pMdc in the cell. There are problems with this technique, however, as putting the gene under control of such a promoter could lead to there being too much, or too little of it available, either of which could have deleterious effects upon the cell.

A third option is to use a similar vector containing an inducible promoter to introduce antisense RNA into the cell to silence the translation of pMdc. Whilst this would leave the gene under the control of its endogenous promoter, it is difficult to determine that complete ablation of function has occurred in the case that pMdc is revealed to be non-essential. Alternately it is also possible that if the promoter used is not strongly repressed when uninduced that enough antisense RNA would be produced to have a silencing effect at a time other than that desired.

Other methods for making conditionally lethal mutations exist, such as amber suppression, whereby a mutant copy of the gene is controlled by the

availability of rare tRNA, however such systems are not yet readily available for *H. volcanii*.

## 8.7 Concluding Statement

In conclusion this study indicates that the carotenoid synthesis pathway of *H. volcanii* is vital for the health of the cell and protective against oxidative damage.

## References

## References

## References

- ALLERS, T., BARAK, S., LIDDELL, S., WARDELL, K. & MEVARECH, M. Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in *Haloferax volcanii*. *Appl Environ Microbiol*, 76, 1759-69.
- ALLERS, T. & MEVARECH, M. 2005. Archaeal genetics - the third way. *Nat Rev Genet*, 6, 58-73.
- ALLERS, T. & NGO, H. P. 2003. Genetic analysis of homologous recombination in Archaea: *Haloferax volcanii* as a model organism. *Biochem Soc Trans*, 31, 706-10.
- ALLERS, T., NGO, H. P., MEVARECH, M. & LLOYD, R. G. 2004. Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the *leuB* and *trpA* genes. *Appl Environ Microbiol*, 70, 943-53.
- AUCHTUNG, T. A., TAKACS-VESBACH, C. D. & CAVANAUGH, C. A. 2006. 16S rRNA phylogenetic investigation of the candidate division "Korarchaeota". *Applied and Environmental Microbiology*, 72, 5077-5082.
- AVELDANO, M. I. 1988. Phospholipid species containing long and very long polyenoic fatty-acids remain with rhodopsin after hexane extraction of photoreceptor membranes. *Biochemistry*, 27, 1229-1239.
- BAKALYAR, S. R., MCILWRICK, R. & ROGGENDORF, E. 1977. Solvent selectivity in reversed-phase high-pressure liquid-chromatography. *Journal of Chromatography*, 142, 353-365.
- BAKER, G. C. & COWAN, D. A. 2004. 16 S rDNA primers and the unbiased assessment of thermophile diversity. *Biochemical Society Transactions*, 32, 218-221.
- BAPTESTE, E. & BOUCHER, Y. 2008. Lateral gene transfer challenges principles of microbial systematics. *Trends in Microbiology*, 16, 200-207.
- BAPTESTE, E., BOUCHER, Y., LEIGH, J. & DOOLITTLE, W. F. 2004. Phylogenetic reconstruction and lateral gene transfer. *Trends in Microbiology*, 12, 406-411.
- BARNS, S. M., FUNDYGA, R. E., JEFFRIES, M. W. & PACE, N. R. Remarkable Archaeal Diversity Detected in a Yellowstone National-Park Hot-Spring Environment. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 1609-1613.
- BARTON, N. H., BRIGGS D.E.G., EISEN J. A., GOLDSTEIN D. B., PATEL, N.H., *Evolution* 2007

## References

- BELL, S. D. 2001. Mechanism and regulation of trnascription in Archaea. *Biochemical Society Transactions*, 29, A48.
- BENDICH, A. J. & DRLICA, K. 2000. Prokaryotic and Eukaryotic chromosomes: What's the difference? *Bioessays*, 22, 481-486.
- BERNANDER, R. 2000. Chromosome replication, nucleoid segregation and cell division in Archaea. *Trends in Microbiology*, 8, 278-283.
- BEVERIDGE, T. J. 2001. The structure of bacterial surfaces and its influence on stainability. *Journal of Histotechnology*, 25, 55-60.
- BIDLE, K. A., HANSON, T. E., HOWELL, K. & NANNEN, J. 2007. HMG-CoA reductase is regulated by salinity at the level of stranscription in *Haloferax volcanii*. *Extremophiles*, 11, 49-55.
- BLAUT, M., MULLER, V. & GOTTSCHALK, G. Energetics of Methanogenesis Studied in Vesicular Systems. *Journal of Bioenergetics adn Biomembranes*, 24, 529-546.
- BLOCHL, E., RACHEL, R., BURGGRAF, S., HAFENBRADL, D., JANNASCH, H. W. & STETTER, K. O. 1997. *Pyrolobus fumarii*, gen. and sp. nov., represents a novel group of archaea, extending the upper temperature limit for life to 113 degrees C. *Extremophiles*, 1, 14-21.
- BREWER, G. J. 2010. Copper Toxicity in the General Population. *Clinical Neurophysiology*, 121, 459-460.
- BRITTON, G. 1998. Overview of Carotenoid Biosynthesis. *Carotenoids*, 3, 13-147.
- BROCHIER, C. FORTERRE, P. & GRIBALDO, S. An emerging phylogenetic core of Archaea: phylogenies of transcription and translation machineries converge following addition of new genome sequences. *BMC Evolutionary Biology*, 5, 36.
- BROCHIER-ARMANET, C., BOUSSAU, B, GRIBALDO, S. & FORTERRE, P. 2008. Mesophilic Crenarchaeota: proposal for a third Archaeal phylum, the Thaumarcheota. *Nature Reviews Microbiology*, 6, 245-252.
- CANN, I. K. O., ISHINO, Y. Archaeal DNA replication: Identifying the pieces to solve a puzzle. *Genetics*, 152, 1249-1267.
- CHABAN, B., NG, S. Y. M. & JARRELL, K. F. 2006. Archaeal habitats – from the extreme to the ordinary. *Canadian Journal of Microbiology*, 52, 73-116.
- CRAMER, P. 2004. RNA polymerase II structure: from core to functional

## References

complexes. *Current Opinion in Genetics and Development*, 14, 218-226.

COGDELL, R. J. 1978. Carotenoids in Photosynthesis. *Philosophical Transactions of The Royal Society of London Series B – Biological Sciences*, 248, 569-579.

COPELAND, H. F. 1956. *The classification of lower organisms*, Palo Alto, Pacific Books.

DANON, A. & STOECKEN, W. 1974. Photophosphorylation in *Halobacterium Halobium*. *Proceedings of the National Academy of Sciences of the United States of America*, 71, 1234-1238.

DISCH, A., SCHWENDER, J., MULLER, C., LICHTENTHALER, H. K. & ROHMER, M. 1998. Distribution of the mevalonate and glyceraldehyde phosphate/pyruvate pathways for isoprenoid biosynthesis in unicellular algae and the cyanobacterium *Synechocystis* PCC 6714. *Biochemical Journal*, 333, 381-388.

DESMOND, E., BROCHIER-ARMANET, C. & GRIBALDO, S. Phylogenomics of the archaeal flagellum: rare horizontal gene transfer in a unique motility structure. *BMC Evolutionary Biology*, 7, 106.

DOOLITTLE, W. F. 2010. The attempt on the life of the Tree of Life: Science, Philosophy and Politics. *Biology and Philosophy*, 25, 455-473.

DOOLITTLE, W. F. & BAPTESTE, E. 2007. Pattern Pluralism and the Tree of Life hypothesis. *Proceedings of the National Academy of Sciences of the United States*, 104, 2043-2049.

DRIDI, B., FARDEU, M. L., OLLIVIER, B., RAOULT, D. & DRANCOURT, M. 2011. The antimicrobial pattern of cultured human methanogens reflects the unique phylogenetic position of Archaea. *Journal of Antimicrobial Chemotherapy*, 66, 2038-2044.

DUNDAS, I. D. & LARSEN, H. 1963. A Study on Killing by Light of Photosensitized Cells of *Halobacterium Salinarium*. *Archiv Fur Mikrobiologie*, 46, 19.

ECKBURG, P. B., LEPP, P. W. & RELMAN, D. A. 2003. Archaea and their potential role in human disease. *Infection and Immunity*, 71, 591-596.

FALB, M., MUELLER, K., KOENIGSMAIER, L., OBERWINKLER, T., HORN, P., VON GRONAU, S., GONZALEZ, O., PFEIFFER, F., BORNBERG-BAUER, E. & OESTERHELT, D. 2008. Metabolism of halophilic Archaea. *Extremophiles*, 12, 177-196.

FORTERRE, P., BROCHIER, C. & PHILIPPE, H. 2002. Evolution of the

## References

Archaea. *Theoretical Population Biology*, 61, 409-422.

FOX, G.E., PECHMAN, K. R. & WOESE, C. R. 1977. Comparative Cataloging of 16S Ribosomal Ribonucleic Acid – Molecular Approach to Procaryotic Systematics. *International Journal of Systematic Bacteriology*, 27, 44-57.

GALANO, A. & FRANCISCO-MARQUEZ, M. 2009. Reactions of OOH Radical with beta-Carotene, Lycopene, and Torulene: Hydrogen Atom Transfer and Adduct Formation Mechanisms. *Journal of Physical Chemistry*, 113, 11338-11345

GLAUERT, H. P., CALFEE-MASON, K., STEMM, D. N., THARAPPEL, J. C. & SPEAR, B. T. 2010. Dietary antioxidants in the prevention of hepatocarcinogenesis: A review. *Molecular Nutrition & Food Research*, 54, 875-896.

GREGOR, D. & PFEIFER, F. 2005. In vivo analyses of constitutive and regulated promoters in halophilic archaea. *Microbiology*, 151, 25-33.

GRIBALDO, S. BROCHIER, C. 2009. Phylogeny of Prokaryotes: does it exist and why should we care? *Research in Microbiology*, 160, 513-521.

GROCHOWSKI, L. L., XU, H. M. & WHITE, R. H. 2006. *Methanocaldococcus jannaschii* uses a modified mevalonate pathway for biosynthesis of isopentenyl diphosphate. *Journal of Bacteriology*, 188, 3192-3198.

GRUNEWALD, K., HIRSCHBERG, J. & HAGEN, C. 2001. Ketocarotenoid biosynthesis outside of plastids in the unicellular green alga *Haematococcus pluvialis*. *J Biol Chem*, 276, 6023-9.

HAGIWARA, T., YASUNO, T., FUNAYAM, K. & SUZUKI, S. 1998. Determination of lycopene, alpha-carotene and beta-carotene in serum by liquid chromatography atmospheric pressure chemical ionization mass spectrometry with selected-ion monitoring. *Journal of Chromatography*, 708, 67-73.

HARPER, J. T., WAANDERS, E. & KEELING, P. J. 2005. On the monophyly of chromalveolates using a six-protein phylogeny of eukaryotes. *Int J Syst Evol Microbiol*, 55, 487-96.

HE, Y. Y. & HADER, D. P. 2002. Reactive oxygen species and UV-B: effect on cyanobacteria. *Photochemical & Photobiological Sciences*, 1, 729-736.

HEBER, D. & LU Q. Y. 2001. Overview of mechanisms of action of lycopene. *Experimental Biology and Medicine*, 10, 920-923.

HU, A., JIAO, N. & ZHANG, C. L. 2011. Community Structure and Function



## References

of Planktonic Crenarchaeota: Changes with Depth in the South China Sea. *Microbial Ecology*, 62, 549-563.

HUBBARD, K. E., NISHIMURA, N., HITOMI, K., GETZOFF, E. D., & SCHROEDER, J. I. 2010. Early abscisic acid signal transduction mechanisms: newly discovered components and newly emerging questions. *Genes and Development*, 24, 1695-1708.

HUBER, H., HOHN, M. J., RACHEL, R., FUCHS, T., WIMMER, V. C. & STETTER, K. O. 2002. A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature*, 417, 63-67.

HUBER, H., HOHN, M. J., STETTER, K. O. & RACHEL, R. 2003. The phylum Nanoarchaeota: Present knowledge and future perspectives of a unique form of life. *Research in Microbiology*, 154, 165-171.

JAHN, U., GALLENBERGER, M., PAPER, W., JUNGLAS, B., EISENREICH, W., STETTER, K. O., RACHEL, R., HUBER, H. 2008. *Nanoarchaeum equitans* and *Ignicoccus hospitalis*: New insights into a unique, intimate association of two Archaea. *Journal of Bacteriology*, 190, 1743-1750.

KARNER, M. B., DELONG, E. F. & KARL, D. M. 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature*, 409, 507-510.

KASHEFI, K. & LOVELY, D. R. 2003. Extending the upper temperature limit of life. *Science*, 301, 934.

KIM, D. S., KIM, J. B., GOH, E. J., KIM, W. J., KIM, S. H., SEO, Y. W., JANG, C. S. & KANG, S. Y. 2011. Antioxidant response of Arabidopsis plants to gamma irradiation: Genome-wide expression profiling of the ROS scavenging and signal transduction pathways. *Journal of Plant Physiology*, 168, 1960-1971.

KLASSEN, J. L. 2010. Phylogenetic and evolutionary patterns in microbial carotenoid biosynthesis are revealed by comparative genomics. *Plos One*, 5, 6.

KLEIN, C., GARCIA-RIZO, C., BISLE, B., SCHEFFER, B., ZISCHKA, H., PFEIFFER, F., SIEDLER, F. & OESTERHELT, D. 2005. The membrane proteome of Halobacterium salinarum. *Proteomics*, 5, 180-97.

KEELING, P. J., CHARLEBOIS, R. L. & DOOLITTLE, W. F. 1994. Archaeobacterial genomes: Eubacterial form and Eukaryotic content. *Current Opinion in Genetics and Development*, 4, 816-822.

KULTZ, D. 2003. Evolution of the cellular stress proteome: from monophyletic origin to ubiquitous function. *J Exp Biol*, 206, 3119-24.

## References

- LANDIS, G. A. 2001. Searching for life: the case for Halobacteria on Mars. *AIP Conference Proceedings*, 552, 25-28.
- LEPP, P. W., BRINIG, M. M., OUVERNEY, C. C., PALM, K., ARMITAGE, G. C. & RELMAN, D. A. Methanogenic Archaea and Human periodontal disease. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 6176-6181.
- LINNAEUS, C. 1758. Tomus I. Systema naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Editio decima, reformata. Holmiae. *Laurentii Salvii*. 1-824.
- LIU, Y. B., FISKUM, G. & SCHUBERT, D. 2002. Generation of reactive oxygen species by the mitochondrial electron transport chain. *Journal of Neurochemistry*, 80, 780-787.
- LOPEZ, A. B., ALEX, B., YANG, Y., THANNHAUSER, T. W. & LI, L. 2008. Phytoene desaturase is present in a large protein complex in the plastid membrane. *Physiologia Plantarum*, 133, 190-198.
- MATHEWS, M. M. & KRINSKY, N. I. 1965. The relationship between carotenoid pigments and resistance to radiation in non-photosynthetic bacteria. *Photochemistry and Photobiology*, 4, 813-7.
- MATSUMI, R., ATOMI, H., DRIESSEN, A. J. M., VAN DER OOST, J. 2010. Isoprenoid biosynthesis in Archaea – Biochemical and Evolutionary Implications. *Research in Microbiology*, 162, 39-52.
- MEYLAN, W. M. & HOWARD, P. H. 1995. Atom/fragment contribution method for estimating octanol-water partition coefficients. *J Pharm Sci*, 84, 83-92.
- MIMURO, M., KATOH, T. 1991. Carotenoids in Photosynthesis – Absorption, Transfer and Dissipation of Light Energy. *Pure and Applied Chemistry*, 63, 123-130.
- MORAN, N. A. & JARVIK, T. 2010. Lateral Transfer of Genes from Fungi Underlies Carotenoid Production in Aphids. *Science*, 328, 624-627.
- MOORE, K. S., WEHRLI, S., RODER, H., ROGERS, M., FORREST, J. N., MCRIMMON, D. & ZASLOFF, M. 1993. Squalamine – An Aminosterol Antibiotic from the Shark. *Proceedings of the National Academy of Sciences of the United States of America*, 90, 1354-1358.
- MULLAKHANBHAI, M. F. & LARSEN, H. 1975. *Halobacterium volcanii* spec nov a Dead Sea *Halobacterium* with a Moderate Salt Requirement. *Archives of Microbiology*, 104, 207-214.

## References

- OLSEN, G. J. 1994. Microbial Ecology – Archaea, Archaea, Everywhere. *Nature*, 371, 657-658.
- OLSEN, G. J., LANE, D. J., GIOVANNONI, S. J., PACE, N. R. & STAHL, D. A. 1986. Microbial Ecology and Evolution – A Ribosomal RNA Approach. *Annual Review of Microbiology*, 40, 337-365.
- OREN, A. 2002. Diversity of halophilic microorganisms: Environments, phylogeny, physiology and applications. *Journal of Industrial Microbiology and Biotechnology*, 28, 56-63.
- OREN, A., GAVRIELI, I., GAVRIELI, J., KOHEN, M., LATI, J. & AHARONI, M. 2005. Microbial Communities in the Dead Sea – Past, present and future. *Cellular Origin and Life in Extreme Habitats and Astrobiology*, 9, 27-39.
- PAEZ, P. L., BECERRA, M. C. & ALBESA, I. 2011. Comparison of Macromolecular Oxidation by Reactive Oxygen Species in Three Bacterial Genera Exposed to Different Antibiotics. *Cell Biochemistry and Biophysics*, 61, 467-472.
- PALOZZA, P., COLANGELO, M., SIMONE, R., CATALANO, A., BONINSEGNA, A., LANZA, P., MONEGO, G. & RANELLETTI, F. O. 2010. Lycopene induces cell growth inhibition by altering mevalonate pathway and Ras signalling in cancer cell lines. *Carcinogenesis*.
- PEDONE, E., BARTOLUCCI, S. & FIORENTINO, G. 2004. Sensing and adapting to environmental stress: The Archaeal tactic. *Frontiers in Bioscience*, 9, 2909-2926.
- REED, A. J., LUTZ, R. A. & VETRIANI, C. 2006. Vertical distribution and diversity of bacteria and archaea in sulfide and methane-rich cold seep sediments located at the base of the Florida Escarpment. *Extremophiles*, 10, 199-211.
- REEVE, J. N. 1999. Archaeobacteria then ... Archaea now (are there really no Archaeal pathogens?) *Journal of Bacteriology*, 12, 3613-3617.
- RONNEKLEIV, M. & LIAAENJENSEN, S. 1995. Bacterial carotenoids 53, C-50-carotenoids 23; Carotenoids of *Haloferax volcanii* versus other halophilic bacteria. *Biochemical Systematics and Ecology*, 23, 627-634.
- SAMUEL, B. S., HANSENS, E. E., MANCHESTER, J. K., COUTINHO, P. M., HENRISSAT, B., RULTON, R., LATREILLE, P., KIM, K., WILSON, R. K. & GORDON, J. I. 2007. Genomic and metabolic adaptations of *Methanobrevibacter smithii* to the human gut. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 10643-10648.

## References

- SANDMANN, G. 2009. Evolution of carotene desaturation: The complication of a simple pathway. *Archives of Biochemistry and Biophysics*, 483, 169-174.
- SANDMANN, G., TAKAICHI, S., FRASER, P. D. 2008. C-35-apocarotenoids in the yellow mutant *Neurospora crassa* YLO. *Phytochemistry*, 69, 2886-2890.
- SAPP, J. The Prokaryote-Eukaryote dichotomy: Meanings and mythology. *Microbiology and Molecular Biology Reviews*, 69, 292.
- SARUHASHI, S. HAMADA, K., HORIIKE, T. & SHINOZAWA, T. 2007. Determination of whole Prokaryotic phylogeny by the development of a random extraction method. *Gene*, 392, 157-163.
- SAUNDERS, N. F. W., THOMAS, T., CUMRI, P. M. G., MATTICK, J. S., KUCZEK, E., SLADE, R., DAVIS, J., FRANZMANN, P. D., BOONE, D., RUSTERHOLTZ, K., FELDMAN, R., GATES, C., BENCH, S., SOWERS, K., KADNER, K., AERTS, A., DEHAL, P., DETTER, C., GLAVINA, T., LUCAS, S., RICHARDSON, P., LARIMER, F., HAUSER, L., LAND, M. & CAVICCHIOLI, R. 2003. Mechanisms of thermal adaptation revealed from the genomes of the Antartic Archaea *Methanogenium frigidum* and *Methanococcoides butonii*. *Genome Research*, 13, 1580-1588.
- SCHWEITZER, J. W. & FRIEDHOFF, A. J. 1966. The metabolism of alpha-14C-3,4-dimethoxyphenethylamine. *Biochem Pharmacol*, 15, 2097-103.
- SEIFERT, J. L. & FOX, G. E. 1998. Phylogenetic mapping of bacterial morphology. *Microbiology UK*, 144, 2803-2808.
- SIEIRO, C., POZA, M., DE MIGUEL, T. & VILLA, T. G. 2003. Genetic basis of microbial carotenogenesis. *Int Microbiol*, 6, 11-6.
- SIMPSON, A. G. B. & ROGER, A. J. 2004. Protein phylogenies robustly resolve the deep-level relationships within Euglenozoa. *Molecular Phylogenetics and Evolution*, 30, 201-212.
- SLOMOVIC, S., PORTNOY, V. & SCHUSTER, G. 2008. Detection and Characterization of Polyadenylated RNA in Eukarya, Bacteria, Archaea and Organelles. *RNA Turnover in Bacteria, Archaea and Organelles*, 447, 501-520.
- SMIT, A. & MUSHEGIAN, A. 2000. Biosynthesis of isoprenoids via mevalonate in Archaea: the lost pathway. *Genome Res*, 10, 1468-84.
- SOSULSKI, F. W. & GADAN, H. M. 1988. Variations in lipid-composition among chickpea cultivars. *Journal of the American Oil Chemists Society*, 65, 369-372.

## References

- TAKANO, E., GRAMAJO, H. C., STRAUCH, E., ANDERS, N., WHITE, J. & BIBB, M. J. 1992. Transcriptional regulation of the REDD transcriptional activator gene accounts for the growth-phase-dependent production of the antibiotic undecylprodigiosin in *Streptomyces coelicolor* A3(2). *Molecular Microbiology*, 6, 2797-2804.
- TIAN, B., XU, Z. J., SUN, Z. T., Lin J. & HUA, Y. J. 2007. Evaluation of the antioxidant effects of carotenoids from *Deinococcus radiodurans* through targeted mutagenesis, chemiluminescence, and DNA damage analyses . *Biochimica et Biophysica Acta-General Subjects*, 6, 902-911.
- VALENZUELA-ENCINAS, C., NEREIA-GONZALEZ, I., ALCANTARA-HERMANDEZ, R. J., ENRIQUEZ-ARAGON, J. A., ESTRADA-ALVARADO, I., HERMANDEZ-RODRIGUEZ, C., DENDOOVEN, L. & MARSCH, R. 2008. Phylogenetic analysis of the archaeal community in an alkaline-saline soil of the former lake Texcoco (Mexico). *Extremophiles*, 12, 247-254.
- WATERS, E., HOHN, M. J., AHEL, I., GRAHAM, D. E., ADAMS, M. D., BARNSTEAD, M., BEESON, K. Y., BIBBS, L., BOLANOS, R., KELLER, M., KRETZ, K., LIN, X. Y., MATHUR, E., NI, J. W., PODAR, M., RICHARDSON, T., SUTTON, G. G., SIMON, M., SOLL, D., STETTER, K. O., SHORT, J. M. & NOORDEWIER, M. 2003. The genome of *Nanoarchaeum equitans*: Insights into an early archaeal evolution and derived parasitism. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 12984-12988.
- WENDOLOSKI, D., FERRER, C. & DYALL-SMITH, M. L. 2001. A new simvastatin (mevinolin)-resistance marker from *Haloarcula hispanica* and a new *Haloferax volcanii* strain cured of plasmid pHV2. *Microbiology UK*, 157, 959-964.
- WHEELIS, M. L., KANDLER, O. & WOESE, C. R. 1992. On the Nature of Global Classification. *Proceedings of the National Academy of Sciences of the United States of America*, 89, 2930-2934.
- WHITTAKER, R. H. 1969 New Concepts of Kingdoms of Organisms. *Science*, 163, 150.
- WIEMER, A. J., HOHL, R. J. & WIEMER, D. F. 2009 The Intermediate Enzymes of Isoprenoid Metabolism as Anticancer Targets. *Anti-Cancer Agents in Medicinal Chemistry*, 9, 526-542.
- WILSON, D. N. 2011. On the specificity of antibiotics targetting the large ribosomal subunit. *Antimicrobial Therapeutics Reviews: Antibiotics that Target the Ribosome*. 1241, 1-16.
- WILSON, H. L., ALDRICH, H. C. & MAUPIN-FURLOW, J. Halophilic 20S

## References

- proteasomes of the archaeon *Haloferax volcanii* Purification, characterization and gene sequence analysis. *Journal of Bacteriology*, 181, 5814-5824.
- WOESE, C. 1998. A manifesto for microbial genomics. *Current Biology*, 8, 781-783.
- WOESE, C. R. & FOX, G. E. 1977. Phylogenetic Structure of Prokaryotic Domain – Primary Kingdoms. *Proceedings of the National Academy of Sciences of the United States of America*, 74, 5088-5090.
- WOESE, C. R., KANDLER, O. & WHEELIS, M. L. 1990. Towards a Natural System of Organisms – Proposal for the Domains Archaea, Bacteria and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America*, 87, 4576-4579.
- WOLF, Y. I., ROGOZIN, I. B., GRISHIN, N. V. & KOONIN, E. V. 2002. Genome trees and the Tree of Life. *Trends in Genetics*, 18, 472-479.
- YACHAI, M. 2009. Carotenoid production by Halophilic Archaea and its Applications.
- YATSUNAMI, R., TAKAICHI, S. & NAKUMURA, S. 2004. Isolation of *crtI* homolog from extremely halophilic archaeon *Haloarcula japonica* strain TR-1. *Nucleic Acids Symposium Series*, 48, 193-4.
- ZHOU, G., KOWALCZYK, D., HUMBAR, M. A., ROHATGI, S. & MAUPIN-FURLOW, J. A. 2008. Proteasomal Components Required for Cell Growth and Stress Responses in the Haloarchaeon *Haloferax Volcanii*. *Journal of Bacteriology*, 190, 8096-8105.
- ZHU, Y. H., JIANG, J. G., YAN, Y. & CHEN, X. W. 2005 Isolation and characterization of phytoene desaturation cDNA involved in the beta-carotene biosynthetic pathway in *Dunaliella salina*. *Journal of Agricultural and Food Chemistry*, 53, 5593-5597.